بسم الله الرحمن الرحيم



جامعة اليرموك
كلية التربية
قسم علم النفس الإرشادي والتربوي

أثر زمن الاختبار على دقة تقديرات معالم الفقرات والأفراد وفق نماذج نظرية استجابة الفقرة

إعداد
هيثم عبدالله عايد منايصة

بإشراف
الأستاذ الدكتور
أحمد سليمان عودة

حقل التخصص- القياس والتقويم
الفصل ألأول 2012 /2013

أثر زمن الاختبار على دقة تقديرات معالم الفقرات والأفراد وفق نماذج نظرية استجابة الفقرة

The effect of test time on the estimations accuracy of items and persons parameters based on the models of Item Response Theory (IRT)

إعداد

هيثم عبدالله عايد منايصة

بكالوريوس: رياضيات – جامعة اليرموك 1983

دبلوم: أساليب تدريس الرياضيات– جامعة اليرموك 1983

ماجستير: إحصاء– جامعة اليرموك 1987

قدمت هذه الأطروحة استكمالا لمتطلبات درجة دكتوراه الفلسفة في تخصص القياس والتقويم التربوي في جامعة اليرموك–اربد،الأردن

وافق عليها

أ.د. أحمد سليمان عودة، .... مشرفا ورئيسا

أستاذ في القياس والتقويم، جامعة اليرموك

أ.د. أمل عبدالله خصاونة.. عضوا

أستاذ في مناهج وأساليب الرياضيات، جامعة اليرموك

أ.د. أحمد يوسف قواسمة، .... عضوا

أستاذ في القياس والتقويم، جامعة اليرموك

د. نضال كمال الشريفين، .... عضوا

أستاذ مشارك في القياس والتقويم، جامعة اليرموك

د. محمد منيزل عليمات، .... عضوا

أستاذ مشارك في القياس والتقويم، جامعة آل البيت

تاريخ مناقشة الأطروحة 26 /2012/12 م

الإهــداء

إلى روح أبي وأمي وأخي عمر.

إلى زوجتي ورفيقة عمري وأبنائي الذين تحملوا المشقة معي.

إلى جميع أخواني وأخواتي وأصدقائي نعم الأخوة والأصدقاء.

اهدي جهدي.

الباحث

شكر وتقدير

بعد أن من الله العلي القدير علي الانتهاء من هذا العمل، لأتوجه إليه عز وجل بالحمد والشكر الذي أعانني على المواصلة والاستمرار، ومن ثم أتقدم بالشكر والعرفان لكل من ساهم في الإشراف عليه منذ أن كان فكرة حتى أصبح حقيقة . وأخص بالذكر الأستاذ الدكتور أحمد سليمان عودة الذي اشرف على هذا البحث وواكب خطواته الواحدة تلو الأخرى مرشدا وموجها ومقوما. ولن أنسى كلا من الأستاذ الدكتورة أمل عبدالله خصاونة، والدكتور نضال كمال الشريفين، والأستاذ الدكتور أحمد يوسف قواسمة، والدكتور محمد منيزل عليمات الذين كان لملاحظاتهم الدقيقة والقيمة أكبر الأثر في بلورة أفكار هذا البحث، فلهم التقدير كله والشكر العميق.

كما أقدم شكري وتقديري لكل من ساعد في تنفيذ إجراءات هذا البحث من أساتذة ومحللين ومدرسين ومديري ومديرات مدارس، وأخوة أوفياء.

وأخيرا أقدم شكري العميق واحترامي للأهل الكرام زوجتي وأبنائي وأخواني وأخواتي الأحباء.

الباحث

المحتوى

قائمة الجداول

الجدول **الصفحة**

قائمة الأشكال

الشكل الصفحة

الملخص

منايصة، هيثم عبدالله، أثر زمن الاختبار على دقة تقديرات معالم الفقرات والأفراد وفق نماذج نظرية استجابة الفقرة. أطروحة بجامعة اليرموك. 2012 (المشرف أ.د. احمد سليمان عودة)

هدفت هذه الدراسة إلى استقصاء أثر زمن الاختبار على دقة تقديرات معالم الفقرات والأفراد وفق نماذج نظرية استجابة الفقرة. ولتحقيق هدف الدراسة، تم بناء اختبار تحصيلي من نوع الاختيار من متعدد في الرياضيات لطلبة الصف التاسع الأساسي، وقد تكوَّن الاختبار بصورته النهائية من (30) فقرة. وشُكِّلت ثلاث صور للاختبار تختلف فقط في الزمن المُعطى للطلبة للإجابة، فأُعطي الطلبة (45) دقيقة للإجابة عن فقرات الاختبار بالصورة الأولى، و(60) دقيقة للإجابة عن فقرات الاختبار بالصورة الثانية، و(75) دقيقة للإجابة عن فقرات الاختبار بالصورة الثالثة. وطُبقت صور الاختبار على عينة مكونة من (1500) طالب وطالبة بواقع (500) طالب وطالبة لكل صورة من صور الاختبار، ثم حُللت الاستجابات لكل صورة بشكل مستقل باستخدام برنامج(BILOG-MG).

وأظهرت النتائج وجود فروق ذات دلالة إحصائية (α=0.05) بين متوسطات الأخطاء المعيارية في تقدير معالم الصعوبة للفقرات، حيث كانت التقديرات لمعالم صعوبة فقرات الصورة الأولى للاختبار أكثر دقة منها لمعالم صعوبة فقرات الصورة الثانية والثالثة، وكانت التقديرات لمعالم صعوبة فقرات الصورة الثالثة للاختبار أكثر دقة منها لمعالم صعوبة فقرات الصورة الثانية للاختبار وفق النموذج الأحادي المعلمة (نموذج راش). ووفق النموذج الثنائي المعالم، كانت التقديرات لمعالم صعوبة فقرات الصورة الأولى للاختبار أكثر دقة منها لمعالم الصعوبة لفقرات الصورة الثانية والثالثة، وكانت التقديرات لمعالم صعوبة فقرات الصورة الثانية للاختبار أكثر دقة

منها لمعالم الصعوبة لفقرات الصورة الثالثة. كما أظهرت النتائج وجود فروق ذات دلالة إحصائية (α=0.05) بين متوسطات الأخطاء المعيارية في تقدير معالم التمييز للفقرات، حيث كانت التقديرات لمعالم تمييز فقرات الصورة الثانية أكثر دقة منها لمعالم تمييز فقرات الصورة الأولى، كما أشارت إلى أن دقة تقديرات معالم التمييز للصورة الثالثة أكثر دقة منها لتقديرات معالم التمييز للصورة الأولى للاختبار، مع عدم وجود دلالة إحصائية بين متوسطات الأخطاء المعيارية في تقديرات معالم التمييز للفقرات بين الصورتين الثانية والثالثة من الاختبار، وذلك وفق النموذج الثنائي المعالم. وأظهرت النتائج أيضا وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (α =0.05) بين متوسطات الأخطاء المعيارية في تقديرات معالم القدرة للأفراد، حيث كانت التقديرات أكثر دقة لقدرات أفراد الصورة الأولى للاختبار منها لتقديرات قدرات أفراد الصورة الثانية والثالثة له، وكانت التقديرات لقدرات أفراد الصورة الثالثة أكثر دقة منها للصورة الثانية وفق النموذج الأحادي المعلمة (نموذج راش). ووفق النموذج الثنائي المعالم كانت التقديرات أكثر دقة لقدرات أفراد الصورة الأولى للاختبار منها لتقديرات قدرات أفراد الصورة الثانية والثالثة له، مع عدم وجود دلالة إحصائية بين متوسطات الأخطاء المعيارية لتقديرات قدرات أفراد الصورتين الثانية والثالثة من الاختبار.

وبالإضافة الى ذلك، أعطت الصورتان الثانية والثالثة أكبر قدر من المعلومات عند مستويات القدرة المنخفضة، بينما أعطت الصورة الأولى أكبر قدر من المعلومات عند مستويات القدرة المتوسطة وفق النموذج الأحادي المعلمة. وأعطت الصورة الأولى أكبر قدر من المعلومات عند مستويات القدرة المتوسطة فقط وفق النموذج الثنائي المعالم. وأوصت الدراسة باستخدام النموذج الثلاثي المعالم في التحليل، واستخدام طريقة تحليل ثانية حيث يتم فيها استبعاد الفقرات غير المطابقة

للنماذج، ثم إجراء التحليل مرة أخرى. (الكلمات المفتاحية: اختبار الاختيار من متعـــدد، صــــعوبة الفقرة، تمييز الفقرة، قدرة الفرد، دالة المعلومات للفقرة، دالة المعلومات للاختبار، زمن الاختبار).

الفصل الأول

خلفية الدراسة والأدب النظري

مقدمة

تستخدم نظرية استجابة الفقرة في تحليل بيانات الاختبارات بشكل واسع لما تتميز به عن النظرية التقليدية في القياس بخصائص عدة، أهمهـــا: أن تقـــديرات الخـــصائص الـــسيكومترية للفقرات، مثل الصعوبة والتمييز مستقلة عن عينة الأفراد التي استخدمت في تقدير هذه الخصائص (Person – Free)، والميزة الثانية أن تقدير قدرة الفرد مستقل عن عينة الفقرات التي تطبق عليه (Item – Free). والخاصية الثالثة هي أن النظرية توفر مؤشرا إحصائيا لتقدير درجة الدقة فـــي قياس القدرة لكل فرد (Hambleton & Swaminathan, 1985).

ومن الاختلافات الرئيسية بين نماذج استجابة الفقرة هي عدد الأبعاد التـــي يفتـــرض أن تكون مسئولة عن أداء المفحوصين على الفقرات المختلفة، وبالتالي يمكن تصنيف نماذج نظريـــة استجابة الفقرة إما إلى نماذج أحادية البعد أو متعددة الأبعاد. (علام، 2005)

ومن النماذج الأحادية البعد (والذي سوف تقتصر الدراسة على هذا النوع من النمـــاذج): النموذج الأحادي المعلمة والنموذج الثنائي المعلمة والنموذج الثلاثي المعلمة. أما النماذج المتعددة الأبعاد فهي: النوع الأول النماذج الكشفية؛ وعادة ما تسمى بالنماذج التعويضية جزئيا؛ ويقصد بها تلك النماذج التي تسمح للأفراد الضعاف في سمة معينة أو أكثر تعويض ضعفهم تعويضا جزئيا؛ أي تسمح هذه النماذج بتفاعل السمات الكامنة أو القدرات متعددة الأبعاد بحيث أن نقص إحـــداها لدى الفرد المفحوص يمكن أن يعوض بقدراته الأخرى المرتفعة؛ أي أن القدرة المرتفعة على أحد

الأبعاد يمكن أن تعوض جزئيا القدرة المنخفضة على بعد آخر من حيث التوصل إلى استجابة صحيحة على الفقرة؛ أي إمكانية حدوث تعويض بين القدرات المطلوبة للاستجابة على فقرة اختباريه. والنوع الثاني النماذج التوكيدية؛ ويمكن تسميتها بالنماذج غير التعويضية جزئيا؛ ويقصد بها تلك النماذج التي تتطلب مستوى معين من كل قدرة بحيث يكون كافيا للتوصل إلى الإجابة الصحيحة عن الفقرة الاختبارية؛ وهذا يعني أن نقص إحدى القدرات لا يعوض تعويضا كاملا بالقدرات الأخرى المرتفعة؛ أي أن زيادة القدرة في أحد الأبعاد لا يتغلب على النقص في بعد آخر.

ويمكن تصنيف النماذج التعويضية الجزئية إلى ثلاث نماذج رئيسية:

1- نموذج راش اللوجستي متعدد الأبعاد (Multidimensional Rash Logistic Model)

2- النموذج اللوجستي ثنائي المعلمة متعدد الأبعاد (Multidimensional Two Parameter Logistic Model)

3- النموذج اللوجستي ثلاثي المعلمة متعدد الأبعاد (Multidimensional Three Parameter Logistic Model)

ويمكن تصنيف النماذج التوكيدية الغير تعويضية إلى نموذجين رئيسيين وفق التصنيف الآتي:

1- النموذج المضاعف ثلاثي المعلمة (The Multiplicative Three Parameter Model)

2- نموذج المكونات المتعددة للسمات الكامنة (Multicomponent Latent Trait Model)

وتتطلب نماذج نظرية استجابة الفقرة (أحادية البعد) تحقق الافتراضات الآتية: افتــراض أحادية البعد (Unidimensionality)، أي أن هناك قدرة واحدة تفسر أداء الفرد في الاختبــار. وافتراض الاستقلال الموضعي (Local Item Independence)، ويقصد بهــذا الافتــراض أن تكون استجابات الفرد عن الفقرات المختلفة في الاختبار مستقلة استقلالا إحصائيا عنــد مــستوى معين من القدرة. وافتراض وتيرية منحنى خصائص الفقــرة (Item Characteristic Curve)، ويقصد بهذا الافتراض أنه يمكن وصف العلاقة بين احتمال الإجابة الصحيحة للمفحــوص عــن الفقرة، وقدرته التي يتم قياسها بالاختبار من خلال اقتران تراكمي صاعد يعرف باســم منحنــى خصائص الفقرة، أو دالة الإستجابة للفقرة، حيث يوفر هذا المنحنى احتمالات إجابة المفحوصــين عن الفقرة في مستويات القدرة المختلفة إجابة صحيحة. وكون المنحنى تراكميا صاعدا، فإنه يشير بوضوح إلى أن احتمال إجابة الفقرة إجابة صحيحة يزداد بازدياد قدرة المفحوص. وفــي العــادة توصف هذه المنحنيات في نماذج الاختبارات المصممة لقياس سمة واحدة أحادية البعــد بدلالــة معلمة واحدة صعوبة الفقرة، أو معلمتين صعوبة الفقرة وتمييزها، أو ثــلاث معــالم الــصعوبة والتمييز والتخمين، حيث يرجع الاختلاف الرئيس بين نماذج السمات المستخدمة إلى اختلاف شكل منحنى خصائص الفقرة. وافتراض التحرر من السرعة في الأداء (speedeness)، حيث تفترض نماذج النظرية الحديثة أن عامل السرعة لا يقوم بدور في الإجابة عن فقرات الاختبار؛ بمعنى أن سبب فشل المفحوصين في الإجابة عن فقرات الاختبار بشكل صحيح هو القدرة المحدودة لــديهم (Limited Ability)، وليس بسبب تأثير عامل السرعة على إجاباتهم عــن فقــرات الاختبــار (Hambleton & Swaminathan, 1985).

إن كلتا النظريتين، النظرية التقليدية في القياس ونظرية استجابة الفقـرة تؤكـدان علـى افتراض التحرر من السرعة. وإذا ما تم انتهاك هذا الافتراض يكون الاختبـار سـريعا؛ أي أن المفحوصين لا يستطيعون الإجابة عن فقرات الاختبار جميعها ضمن الوقت المسموح فيه، فإما أن يتركوا بعض الفقرات دون إجابة، أو يخمنوا إجاباتها وخصوصا تلك الفقرات التي تقع في نهايـة الاختبار (والتخمين هنا ليس بمستوى التخمين العشوائي الذي تفترضة النظريتين)، وبالتالي فـإن سرعة الاختبار تؤدي إلى تقديرات غير دقيقة لمعالم الفقـرة، وإحـصائيات الاختبـار، وقـدرة المفحوصين (Douglas, Kim, Habing & Gao, 1998; Oshima, 1994).

ومن المشكلات التي تظهر في إطار هذه النظرية هي صعوبة تحقق بعض الإفتراضات، وركزت الدراسات على تأثير انتهاك افتراضات أحادية البعـد بالدرجـة الأولـى، والاسـتقلال الموضعي التي تنتهك بسهولة، إلا أن الأبحاث التي تناولت أثر انتهاك افتـراض التحـرر مـن السرعة على تقديرات القدرة والفقرة لا زالت متواضعة. وعندما يكون الاختبار سريعا لدرجة ما، فإن أثر فشل المفحوصين في الإجابة عن الفقرات التي لم يتم الوصول إليها، أو الإجابـة عنهـا عشوائيا، كلاهما يقلل من صدق الاختبار وثباته، وبالتالي يحتاج إلى بحث واستقصاء (Oshima, 1994).

ولأغراض هذه الدراسة، تم التركيز على افتراض التحرر من السرعة المتمثـل بعامـل اختلاف زمن الاختبار مع الثبات النسبي أو التكافؤ في المؤشرات الإحصائية للافتراضات الأخرى الأساسية، وتم تناول هذا الافتراض من حيث: المعاني المختلفة لافتراض التحرر من الـسرعة، والتأثيرات الناجمة عن انتهاك هذا الافتراض على كل من إحصائيات الفقرة والاختبـار وقـدرة

المفحوصين، وعلاقة افتراض التحرر من السرعة بافتراض أحادية البعد، كما تم تناول مناقشة آراء الباحثين حول العلاقة بين سرعة الاختبار ووضع حدود زمنية للاختبار.

المعاني المختلفة لافتراض التحرر من السرعة

تفترض كل من النظرية التقليدية في القياس ونظرية استجابة الفقرة أن الاختبار يكون متحررا من السرعة. وأوضح الباحثون نوعين من الاختبارات تبعا لسرعة الاختبار: النوع الأول اختبارات قوة (Power Tests)، حيث ذكر سجنبيك (Schnipke, 1995) أن الاختبار يسمى " اختبار قوة " عندما يكون لدى معظم المفحوصين الوقت الكافي لإنهاء الاختبار. وعرف هامبلتون وسواميناثان (Hambleton & Swaminathan, 1985) بمعنى دقيق أن الاختبار يسمى اختبار قوة، إذا كان مجموع علامات الشخص على الفقرات التي يجيب عنها إجابة غير صحيحة يساوي ببساطة عدد الفقرات التي يجيب عنها بشكل خطأ، وبحيث لا يكون للوقت دور ضاغط على المفحوص في الإجابة. وأن في اختبار" القوة التامة" (اختبار القوة 100%) (Pure Power Test) يتم المحاولة في الفقرات جميعها، ولهذا فإن علامات الاختبار تعتمد بشكل رئيسي على عدد الفقرات التي تم الإجابة عنها بشكل صحيح، والتي تم الإجابة عنها بشكل خطأ، ولا يمتلك هذا الاختبار حدودا زمنية، وفي ضوء هذا التعريف يتم الافتراض بعدم حذف أي فقرة أو إهمالها (Peterson,1993).

أما النوع الآخر من الاختبارات، فهو ما يسمى بإختبارات السرعة، حيث أشار لو وسيرسي (Lu and Sireci, 2007) إلى أن "السرعة" تشير إلى الوضع الذي لا يسمح فيه الوقت المحدد للاختبار المعياري (معايرة الزمن) لعدد كبير من المفحوصين من التفكير في فقرات

الاختبار جميعها. وذكر بيجار (Bejar, 1985, p.1) أن "الاختبار يعد سريعا عندما لا يملك بعض المفحوصين الوقت الكافي لمحاولة الإجابة عن كل فقرة في الاختبار ضمن الوقت المعطى لهم (المسموح به)". وترى مؤسسة الاختبارات التعليمية الأمريكية (Educational Testing Service (ETS)) أن الاختبار يسمى "اختبار سرعة" عندما لا يتمكن بعض المفحوصين من إنهاء الإجابة على فقرات الاختبار جميعها، ويمكن قياس السرعة بنسبة عدد المفحوصين الذين لم يتمكنوا من الوصول إلى جميع الفقرات أو لم يتمكنوا من الوصول إلى 75% من الفقرات (Hecht and Swinford, 1981; Swinford, 1979). وتم تعريف الاختبار "السريع التام" (Pure Speededness) على أنه اختبار ذو فقرات في غاية السهولة، بحيث أن المفحوص لا يعطي إجابة خطأ على أي فقرة، ومع ذلك، فإن الاختبار يحتوي على فقرات كثيرة، بحيث لا يتمكن أحد من إنهائه في الوقت المسموح فيه، وبالتالي فإن علامات المفحوصين تحدد بعدد الفقرات التي تمت محاولة الإجابة عنها (Peterson,1993).

وتوجد فروق عدة بين اختبار القوة واختبار السرعة، ففي اختبار القوة التامة تتمتع فقراته بتدرج في الصعوبة، ولا يوضع حدود زمنية له، والهدف منه قياس مدى دقة إجابة المفحوص عن الفقرات. أما في اختبار السرعة التامة فتتميز فقراته بالسهولة ويتحدد الزمن اللازم له بدقة، والهدف هنا هو قياس مدى سرعة إجابة المفحوصين عن الفقرات. ومن ناحية عملية فإن الاختبار يحتوي على جزء من القوة، وجزء من السرعة، وبالتالي يعتبر اختبار سريع؛ أي أن الاختبار السريع عادة ناتج عن اختبار قوي، ولكن محدد بوقت معين، وهذا هو المطلوب عمليا في معظم الأحيان (Peterson,1993).

أثر انتهاك افتراض التحرر من السرعة

في سياق نظرية استجابة الفقرة يعتمد احتمال الاستجابة الصحيحة للمفحوصين عن الفقرة على أداء المفحوصين وخصائص الفقرة فقط، إضافة إلى ذلك تفترض نظرية استجابة الفقـــرة أن المفحوصين يمتلكون الوقت الكافي للإجابة عن جميع الفقرات في الاختبار، ولكن في الاختبارات السريعة لم تُصمم تلك الاختبارات لتأخذ السرعة بالحسبان، وبالتالي يمكن انتهاك افتراض الوقت الكافي في نظرية استجابة الفقرة عندما يفشل المفحوصون في الإجابة الـــصحيحة عـــن فقـــرات الاختبار، ليس بسبب القدرة المحدودة لديهم، ولكن بسبب الوقت المحدود. إذن الاستجابة السريعة غير الملاحظة أو غير المكتشفة (Undetected) تسمح لعامل الوقت أن يشوه (Contaminate) تقديرات القدرة. وعلاوة على ذلك، فإن المفحوصين غالبا ما يسرعوا في الإجابة عـــن الفقـــرات الأخيرة للاختبار قبيل انتهاء وقت الاختبار، أو يحذفوا تلك الفقرات، أو يكملوا الإجابة عنها بشكل عشوائي، وبالتالي تبدو هذه الفقرات أكثر صعوبة من تلك التي يُجاب عنها عنـــدما تـــدار تحـــت ظروف غير سريعة (Oshima, 1994).

وفي ظل الحرية التي تعطى للطالب بأن يبدأ بالسؤال الذي يريد، فمن غير الممكن الجزم بأن الأسئلة أو الفقرات التي تتعرض لتأثير الوقت هي التي تقع في نهاية ورقة الامتحان، فلكـــل مفحوص طريقتة في التعامل مع التسلسل في الإجابة وإدارة وقت الاختبار، ولذلك فإن من أحـــد معاني السرعة هي: السرعة في زمن التفكير وفي كل العوامل المحددة للإجابة، وتكون خاضـــعة لمبدأ الفروق الفردية.

وإن انتهاك افتراض التحرر من السرعة له نتائج سلبية عند تطبيق نماذج استجابة الفقرة أحادية البعد، ومن أخطرها وأكثرها أهمية تلك المتعلقة بتقدير معالم الفقرة، وتقدير معالم المفحوصين، التي تؤدي بالتالي إلى تهديد صدق الاختبار، وتسبب مشاكل عديدة عند بناء تجمع من الفقرات (Pool of Items) الذي يستخدم لإيجاد أشكال جديدة من الاختبارات و/ أو المحافظة على تدريج للعلامات (Scale of Scoring)، مثال ذلك، عند بناء الاختبارات المتكافئة.

وبينت العديد من الدراسات (Bolt, Cohen & Wollack, 2002; Kingston & Dorans, 1984; Oshima, 1994)، أنه نتيجة لانتهاك افتراض التحرر من السرعة، فإن الفقرات التي توضع في نهاية اختبار معين تميل إلى أن تمتلك خصائص مختلفة عندما توضع نفس الفقرات في مكان متقدم في شكل آخر من أشكال الاختبار نفسه (أي أن نتيجة السرعة فإما أن يتركوا المفحوصون الفقرات الأخيرة دون إجابة أو يخمنوا إجاباتهم وهذا ما يعني أن تلك الفقرات تبدو على أنها صعبة، ولكنها في الحقيقة ليست بهذه الصعوبة لو وضعت في مكان آخر غير نهاية الاختبار)، وبالتالي فإن ذلك يؤدي إلى عدم الاستقلال الموضعي وبالتالي يؤدي إلى عدم دقة تقدير معلمة الصعوبة، وكذلك يؤدي إلى عدم دقة تقديرات معلمات استجابة الفقرة الأخرى، كمعلمة التمييز، ومعلمة التخمين. وغالبا ما يؤدي عدم الدقة في تقديرات معالم الفقرة إلى مشاكل عدة خصوصاً عند بناء برنامج يهدف إلى المحافظة على تجمع من الفقرات (Mroch, 2006).

التحرر من السرعة وأحادية البعد

لقد ظهرت العديد من الآراء ووجهات النظر للباحثين وأصحاب الاختصاص حول افتراضي أحادية البعد والتحرر من السرعة والعلاقة بينهما، فمعظم الباحثين اعتبر أن افتراض أحادية البعد يتضمن افتراض التحرر من السرعة.

وأبرز الآراء التي أيدت شمول افتراض أحادية البعد لافتراض التحرر من السرعة ما ذكره هامبلتون وسواميناثان (Hambleton & Swamint2006, p.5) حيث بينا أنه من النادر الإشارة إلى افتراض التحرر من السرعة؛ لأنه قد يكون متضمنا في افتراض أحادية البعد، فعندما تؤثر السرعة في الأداء على الاختبار، فإن هذا يعني أن هنالك سمتين على الأقل تؤثران في الأداء، هما: سرعة الأداء، والسمة التي يسعى الاختبار لقياسها.

واتفق مروج (Mroch, 2006) مع هامبلتون وسواميناثان (Hambleton & Swaminthan,1985) في قوله إن انتهاك افتراض التحرر من السرعة يؤدي إلى انتهاك افتراض أحادية البعد في نظرية استجابة الفقرة، وذلك لأن سرعة الاختبار تضيف بعدا إحصائيا آخر لما يقيسه الاختبار؛ أي أن هنالك عامل آخر في الاختبارات السريعة يمكن استخراجه، وبالتالي يصبح الاختبار السريع متعدد الأبعاد.

وهذا يشكِّل مشكلة للاختبارات التي تقيس سمة واحدة، ففي نظرية استجابة الفقرة على سبيل المثال، تعد النماذج (النماذج أحادية البعد) عادة لقياس قدرة وحيدة؛ أي أنها تفترض افتراض أحادية البعد (مثال ذلك النموذج أحادي المعلمة (نموذج راش)، والنموذج ثنائي المعالم والنموذج ثلاثي المعالم، وبالتالي بوجود عامل السرعة، فإن نموذج أحادية البعد يمكن أن يصف سمة البعد

الواحد بأنها مركبة من السمة المقصودة وعامل السرعة. ونتيجة انتهاك افتراض أحادية البعد الواحد في نظرية استجابة الفقرة تحدث مشاكل عدة أهمها عدم دقة تقديرات معالم الفقرة (Mroch, 2006).

كما أن انتهاك افتراض التحرر من السرعة يؤدي إلى انتهاك افتراض الاستقلال الموضعي في نظرية استجابة الفقرة، فالسرعة تمهد إلى عدم الاستقلال الموضعي بين الفقرات خصوصاً بين الفقرات التي تقع في نهاية الاختبار على افتراض أنها تركت بدون إجابة، وبالتالي يسبب انتهاك افتراض الاستقلال الموضعي للفقرات مشاكل أهمها تعدد الأبعاد وعدم دقة تقديرات معالم الفقرة، خصوصا وأن بعض الدراسات تعاملت مع الاستقلال الموضعي من خلال افتراض أحادية البعد (Mroch, 2006)، بمعنى أن تحقق افتراض أحادية البعد يحقق تلقائياً افتراض الاستقلال الموضعي. وفي هذا المجال، ظهرت العديد من الآراء ووجهات النظر وأصحاب الاختصاص حول افتراضي أحادية البعد والاستقلال الموضعي والعلاقة بينهما، فمنهم من أعتبر هذين الافتراضين متكافئين، ومنهم من قال خلاف ذلك.

وأبرز الآراء التي قالت في تكافؤ المفهومين ما ذكره هامبلتون وسواميناثان (Hambleton & Swaminthan,1985) حيث بينوا هذا التكافؤ من زاويتين، فلو اعتبرنا أن مجموعة من الفقرات تقيس سمة واحدة، فإن المفحوصين الذين يشتركون في نفس القدرة ستكون استجاباتهم على الفقرات مستقلة إحصائيا، ومن هنا فلو لم تكن الفقرات مستقلة إحصائيا عند مستوى قدرة معين، فهذا يؤكد لنا بأن بعض المفحوصين يتوقع أن تكون علاماتهم أعلى من المفحوصين الآخرين الذين يشتركون في نفس القدرة، ونتيجة لذلك سيكون هنالك أكثر من قدرة

واحدة مسؤولة عن أداء المفحوصين، وهذا انتهاك واضح للافتراض الرئيسي بأن الفقرات تقيس سمة واحدة. والنقطة ألاخرى هو ان افتراض الاستقلال الموضعي يتضمن أن الاستجابات على الفقرات المختلفة مستقلة إحصائيا بالنسبة للمفحوصين عند مستوى قدرة معين، وبذلك لا بد وأن تكون هناك قدرة واحدة فقط هي المسئولة عن الارتباطات بين مجموعة فقرات الاختبار.

ومن الآراء التي قالت بعدم تكافؤ مفهومي أحادية البعد والاستقلال الموضعي ، ذكر ميرا وروبين وسيرسي (Meara, Robin & Sireci, 2000) بأن افتراض أحادية البعد لا يكافئ افتراض الاستقلال الموضعي، موضحين أنه عند مستوى قدرة معين، يمكن أن تكون فقرات الاختبار مستقلة إحصائيا ويقيس هذا الاختبار أكثر من سمة واحدة، وأشاروا بأن أحادية البعد في نماذج نظرية استجابة الفقرة التي تشدد على هذا الافتراض تعرف بناءا على وجود سمة واحدة التي يكون عندها الاستقلال الموضعي محققا.

العلاقة بين سرعة الاختبار ووضع حدود زمنية للاختبار

لا تزال مسألة وضع حدود زمنية للاختبارات تثير جدلا بين معظم الباحثين، فمنهم من يرى ضرورة لذلك إذا كانت سرعة الأداء هي الجزء الرئيس المستهدف من الاختبارات كقياس سرعة شخص في الطباعة. كما أن الوقت المحدد يساعد أيضا في تقدير تكلفة إدارة الاختبارات كتكلفة رسوم الساعة للمراقبين في اختبارات الورقة والقلم، أو تكلفة زمن المراقبة في مراكز الاختبارات المحوسبة لاختبارات تعليم اللغة الإنجليزية كلغة أجنبية (TOEFL) (Bridgeman, McBride & Monaghan, 2004). ويمكن أن يهدد تحديد مدة الاختبار بصورة كبيرة صدق الاختبار، أو يهدد ما صُمم الاختبار لقياسه.

وغالبا ما ترجع سرعة الاختبار إلى تأثير حدود الزمن على علامات المفحوصين عندما تكون حدود وقت الاختبار مقيدة إلى الحد الذي لا يجد عنده معظم المفحوصين الوقت الكافي للتفكير والإجابة عن كل فقرة في الاختبار، وعندها يقال إن الاختبار سريع "Speeded"، وكذلك يسمى الاختبار سريعا إلى مدى معين، إذا كان المفحوصون يأخذون علامات أقل من العلامات التي يأخذونها إذا كان لديهم الوقت الكافي لإنهاء الاختبار، أو أعطوا وقتا إضافيا غير محدد لإنهائه (Bridgeman, McBride & Monaghan, 2004).

إضافة إلى أسباب عملية ضرورية، فإن الدافع الأساسي لوضع حدود زمنية دقيقة محددة يكمن في التأكد من أن المفحوصين جميعهم خضعوا لنفس الظروف المعيارية ومنها زمن الاختبار الموحد، وبالتالي فإن في هذا السياق على الأقل يكون الاختبار عادلا و منصفا (Fair)، لكل مفحوص(Powers & Fowles, 1996) ومن الصعب جداً تفصيل زمن يتناسب مع خصائص كل مفحوص.

وتستخدم معدلات إكمال الاختبار بعض الأحيان لافتراض سرعة الاختبار؛ أي أن علامات المفحوصين في الاختبار تتأثر بالوقت المحدد بشكل واضح، خصوصا إذا كان عدد كبير من المفحوصين يفشلون في إكمال الاختبار، وعندها يكون الاختبار سريعا، ولكن غالبا ما يمنع الوقت المحدد للاختبارات المفحوصين من أن ينهوا الفقرات جميعها، وهذا ما يسمى "سرعة" الاختبار (Bridgeman , Cline & Hessinger, 2003). كما يمكن ملاحظة سرعة الاختبار غالباً عند وضع قيود زمنية على أداء المفحوصين في الاختبار، حيث أشار ريندلر (Rindler, 1979, p.261)، إلى أن "الاختبار يعد أداة سريعة جزئياً عندما يؤثر الوقت المحدد في علامات

المفحوصين"؛ أي أن علامة الشخص تحدد بعدد الفقرات التي يحاول فيها (سرعة الاستجابة) وبدقة الاستجابات المعطاة (قوة الاستجابة). ووجد أن معظم الاختبارات المعيارية تدار بوقت ثابت ومحدد، وأن هذه الحدود تُقر لإتاحة الفرصة لمعظم المفحوصين المحاولة في الفقرات جميعها، أي أن لا يترك المفحوصون أي فقرة محذوفة. وفي هذا السياق عرف نانلي (Nunnally, 1978,p.656) الوقت المحدد المريح "بمقدار الوقت اللازم، ليتمكن معظم المفحوصون أن يكملوا الاختبار في ظل الظروف الطبيعية". ويترك للمفحوصين أن يحدد بنفسه الزمن الذي يستغرقه في الإجابة من الزمن المعطى له، وهنا تظهر حكمة المفحوص في كيفية استثمار الوقت وإدارته في ضوء العوامل الشخصية والظرفية الزمانية منها والمكانية التي تحدد بمجموعها كيفية تصرف المفحوص أثناء الاختبار، فقد يتحول إلى سرعة مع أنه ليس اختبار سرعة والعكس صحيح. ولذلك من الصعب معرفة هذه التفاصيل وكيفية مساهمتها في الإجابة أثناء الاختبار ضمن الزمن المحدد له، وكيفية تأثيرها على خصائص الفقرات وعلى مدى تحقق الافتراضات التي تقوم عليها نظرية القياس. وبالتالي عندما تدار الاختبارات ضمن قيود زمنية محددة، فإن أداء المفحوصين يتأثر بالسرعة، ونتيجة التأثر بالسرعة يؤدي ذلك إلى عدم دقة تقديرات معالم الصعوبة ومعالم التمييز ومعالم التخمين في نماذج نظرية استجابة الفقرة (Oshima, 1994).

مشكلة الدراسة وأسئلتها

يتضح مما سبق أن هناك فرصة دائماً أو غالباً إلى انتهاك افتراض السرعة. وأن انتهاك مثل هذا الافتراض يؤدي إلى تقديرات غير دقيقة فيما يخص قدرات المفحوصين، ومعالم الفقرة وإحصائيات الاختبار.

وقد اختلفت المؤسسات و الدراسات في الاستراتيجيات أو التصاميم المستخدمة في التعامل مع هذا الافتراض، فمؤسسة الاختبارات التعليمية الأمريكية تعتبر أن الاختبار يحقق افتراض السرعة غالبا إذا كان جميع المفحوصين يحاولون الإجابة عن 75% من فقرات الاختبار، وعلى الأقل أن يصل 80% من المفحوصين إلى الفقرة الأخيرة في الاختبار (Hecht and Swinford, 1981; Swinford, 1979).

حاولت الدراسة الحالية أن تكون حقيقية البيانات في تصميمها لفحص هذا الافتراض وتأثيره على خصائص الأفراد والفقرات، وخاصة في تأثيرها على الاختبارات المدرسية التي يجريها المعلمون باستمرار في المباحث المختلفة. فقد تم التحكم بزمن الإجابة بأقل وبأكثر من الزمن المقدر لفترة زمنية جوهرية، دون التطرق للعوامل والمتغيرات الشخصية والظرفية التي قد تعمل على انتهاك افتراض عدم السرعة. واقتصار التعريف الإجرائي على زمن الاختبار، ثم بحث أثره على تقدير القدرة للمفحوصين، وتقدير معالم الفقرة وفقاً لنظرية استجابة الفقرة، وذلك من خلال بيانات حقيقية تتضمن إعداد اختبار حقيقي.

وبالتالي، يمكن تلخيص هدف الدراسة بما يأتي، معرفة تأثير انتهاك افتراض السرعة للاختبار المتمثل بعامل زمن الاختبار على دقة تقديرات معالم الفقرات والأفراد وفق نظرية استجابة الفقرة المختلفة، وذلك من خلال استخدام نموذجين من النماذج اللوجستية لنظرية استجابة الفقرة (الأحادي والثنائي) باستخدام بيانات حقيقية، وبالتحديد فإن مشكلة الدراسة تتلخص في الإجابة عن الأسئلة التالية:

1. هل تختلف دقة تقدير معالم الصعوبة لفقرات اختبار اختيار من متعدد باختلاف الزمن المعطى للاختبار والنموذج المستخدم (أحادي المعلمة، ثنائي المعلمة)؟
2. هل تختلف دقة تقدير معالم التمييز لفقرات اختبار اختيار من متعدد باختلاف الزمن المعطى للاختبار والنموذج المستخدم (ثنائي المعلمة)؟
3. هل تختلف دقة تقدير معالم القدرة لأفراد اختبار اختيار من متعدد باختلاف الزمن المعطى للاختبار والنموذج المستخدم (أحادي المعلمة، ثنائي المعلمة)؟
4. هل تختلف كمية المعلومات للفقرات وللاختبار عند كل مستوى من مستويات القدرة لأفراد اختبار اختيار من متعدد باختلاف الزمن المعطى للاختبار والنموذج المستخدم (أحادي المعلمة، ثنائي المعلمة)؟

متغيرات الدراسة

اشتملت الدراسة الحالية على المتغيرات الآتية:

1- المتغيرات المستقلة: حالة انتهاك افتراض السرعة المتمثل بعامل زمن الاختبار. للاختبار ثلاث صور متكافئة (لها نفس الفقرات) تختلف فقط في زمن الاختبار المُعطى للطلبة للإجابة عن فقرات الاختبار، فأُعطي الطلبة (45) دقيقة للإجابة عن فقرات الاختبار بالصورة الأولى للاختبار، و(60) دقيقة للإجابة عن فقرات الاختبار بالصورة الثانية، و(75) دقيقة للإجابة عن فقرات الاختبار بالصورة الثالثة.

2- المتغيرات التابعة: التقديرات المختلفة لنظرية استجابة الفقرة، في حالة تغيير زمن الاختبار. وهي: معالم قدرات المفحوصين، ومعالم الصعوبة، ومعالم التمييز، ودالة معلومات الاختبار.

التعريفات الاصطلاحية والإجرائية

اختبار الاختيار من متعدد

اختبار كل فقرة من فقراته، لها أربعة بدائل، يعطى الفرد العلامة (1) عند اختيار البديل الصحيح، والعلامة (0) عند اختيار البديل الخطأ.

صعوبة الفقرة

صعوبة الفقرة : هي أحد معالم دالة الاستجابة للفقرة وتعادل قيمة القدرة التي تناضر احتمال 50% للإجابة الصحيحة عندما تكون قيمة التخمين صفرا، وتعادل قيمة القدرة التي تناضر احتمال $(c+1)/2$.

تمييز الفقرة

يستدل على التمييز من خلال ميل المنحنى وليس ميل المماس. التمييز يمكن ملاحظته في ازدياد المنحنى المميز له بالقرب من الوسط، فالميل الحاد يدل على أن التغيير الصغير نسبيا في قدرة الفرد يجعل احتمال الإجابة الصحيحة عن الفقرة يتباين تباينا ملحوظا (تمييز مرتفع) أما الميل البسيط فيدل على أن التغير الكبير نسبيا في قدرة الفرد لا يغير احتمال الإجابة الصحيحة تغيرا ملحوظا (تمييز منخفض)، لذلك فإن تمييز الفقرة يتناسب تناسبا طرديا مع ميل المنحنى المميز لها عند نقطة انقلابه.

قدرة الفرد

من أولويات نظرية استجابة الفقرة تقدير قدرات المفحوصين للمقارنة بينهم، وتستخدم هذه النظرية عدة طرق لتقدير القدرة منها: طريقة الارجحية العظمى (Maximum Likelihood) لتقدير القدرات على مجموعة من الفقرات معلومة الخصائص. وبالتالي فالقدرة هي قيمة يتم تقديرها تعظم استجابات الفرد عن فقرات الاختبار.

دالة المعلومات للفقرة

إن مفهوم المعلومات مثل دالة الاستجابة للفقرة، وهو مفهوم أساس في نظرية استجابة الفقرة(علام، 2005) والمعلومات في القياس النفسي تمثل مقدار الدقة في الشيء(المقيس) المقدر، وقد عرفها فيشر Fisher بأنها بديل للدقة مع أي معلمة تكون مقدرة. وفي سياق نظرية استجابة الفقرة يكون من الممكن أن نحدد كمية المعلومات (أقصى ارتفاع لمنحنى الدالة) عند أي مستوى للفقرة (Θ) يكون أكثر دقة للقياس عند هذا المستوى. وعلى ذلك تظهر أهمية نظرية استجابة الفقرة في قدرتها على تحديد أي الفقرات تقيس المتغير بدرجة أفضل عند مستويات محددة للقدرة، وأن الطريقة الفعالة لوصف الفقرات والاختبارات وانتقاء فقرات الاختبارات ومقارنة الاختبارات تقدم بوساطة نظرية استجابة الفقرة من خلال استخدام دوال المعلومات،ودوال معلومات الفقرة تؤدي دورا مهما في بناء الاختبار وتقويم الفقرة التي يتضح من خلالها إسهام الفقرات المختارة في تقدير القدرة عند النقاط المختلفة لـ (Θ) عبر متصل القدرة.

دالة المعلومات للاختبار

دالة معلومات الاختبار تقدم وسيلة قوية لمصمم الاختبار لأغراض التقويم لخواص الاختبار في ضوء كيف يقدم الاختبار أفضل قياس عبر متصل السمة أو القدرة، فضلا على أنها أداة تساعد في تقييم مدى الدقة المقدمة من خلال الاختبار عند أي مستوى للسمة الكامنة.

وتعد دوال معلومات الفقرات حجر الأساس لدوال معلومات الاختبار، وذلك لأن دالة معلومات الاختبار تعتمد على تجميع إسهامات كل فقرة من فقرات الاختبار من المعلومات عند مستوى القدرة **(Ө)**.

زمن الاختبار

للاختبار ثلاث صور تختلف فقط في طول زمن الاختبار. فأُعطي الطلبة (45) دقيقة للإجابة عن فقرات الاختبار بالصورة الأولى، و(60) دقيقة للإجابة عن فقرات الاختبار بالصورة الثانية، و(75) دقيقة للإجابة عن فقرات الاختبار بالصورة الثالثة.

محددات الدراسة

1- اقتصرت الدراسة على استخدام النموذج اللوجستي الأحادي المعلمة والنموذج اللوجستي ثنائي المعلمة .

2- اهتمت الدراسة بافتراض السرعة (المتمثل بزمن الاختبار) للنماذج أحادية البعد فقط.

الفصل الثاني

الدراسات السابقة

تعددت الدراسات التي بحثت في أثر انتهاك افتراضات نظرية استجابة الفقـرة إلا أن الدراسات التي تناولت أثر انتهاك افتراض التحرر من السرعة لا زالت متواضعة. وفي ضوء هدف الدراسة الحالية المتمثل في دراسة أثر زمن الاختبار على دقة تقديرات معـالم الفقـرة والأفراد وفق نماذج نظرية استجابة الفقرة ، وجد الباحث أن هنالك نـوعين مـن الدراسـات المرتبطة بهذه الدراسة، الأول: الدراسات المتعلقة بتقييم عامل السرعة وأثره على تقدير معالم الفقرات والأفراد في نظرية استجابة الفقرة، والثاني: الدراسات المتعلقة بآثار زمن الاختبـار على أداء المفحوصين.

أولا: الدراسات التي بحثت في تقييم عامل السرعة وأثره علـى تقـديرات معـالم الفقـرات والأفراد في نظرية استجابة الفقرة

ثانيا: الدراسات التي بحثت في أثار زمن الاختبار على أداء المفحوصين

أولا: الدراسات التي بحثت في تقـدير الـسرعة وأثـره علـى تقـديرات معـالم الفقـرات والأفراد في نظرية استجابة الفقرة

هدفت دراسة لبيترسون (Peterson, 1993) إلى مراجعة طرق تقدير سرعة الاختبارات (والتي تسمى بالطرق التقليدية لتقدير السرعة)، حيث تم تصنيفها إلى مجموعتين

رئيسيتين: الطرق التي تقدر السرعة بناء على تطبيق واحد للاختبارات، والطرق التي تقدر السرعة بناء على تطبيق متعدد للاختبارات.

أولا: مجموعة الطرق التي تقدر السرعة بناء على تطبيق واحد للاختبارات، من أهمها:

1. نسبة جلكسن (Gulliksen, 1950) :

نسبة جلكسن= (الانحراف المعياري للفقرات التي لم يتم المحاولة فيها) / (الانحراف المعياري للفقرات التي لم يتم المحاولة فيها + الانحراف المعياري للفقرات التي تم الإجابة عنها بشكل خطأ(أو الفقرات المحذوفة بعد التفكير فيها))

فإذا كان الاختبار غير سريع 100%، فإن الانحراف المعياري للفقرات التي لم يتم المحاولة فيها يساوي (صفر)، وتكون النسبة عندها تساوي (صفر)، وإذا كان الاختبار السريع 100%، يكون الانحراف المعياري للفقرات المجابة خطأ (صفر) والنسبة عندها تساوي (1).

2. **طريقة مؤسسة الاختبارات التربوية الأمريكية (Educational Test Service (ETS)) (Donlon, 1973)** التي تفيد أن الاختبار يكون سريعا إذا كان أقل من 100% من المفحوصين يصلون على الأقل إلى 75% من الفقرات، وأقل من 80% من المفحوصين ينهون الاختبار.

ثانيا: مجموعة الطرق التي تقدر السرعة بناء على تطبيق متعدد للاختبارات، والتي تفيد بأن الاختبار يكون غير سريع إذا بقي ترتيب المفحوصين ثابتا، سواءً أأعطوا الاختبار في ظل ظروف سرعة، أو في ظل ظروف عدم السرعة، ومن أهمها: طريقة كرونباخ

ووارينغتون (Cronbach &Warrington, 1951) حيث استخدما معامل الارتباط تاو (tau) الذي يمثل مربع معامل الارتباط بين العلامات في الاختبار السريع، وعلامات اختبار موازي يدار دون زمن محدد.

ومن الأبحاث التي هدفت إلى دراسة أثر انتهاك افتراض التحرر من السرعة على تقديرات القدرة، ومعالم الفقرة في نظرية استجابة الفقرة، دراسة أوشيما (Oshima, 1994) حيث استخدمت فيها أكثر من عامل من عوامل انتهاك افتراض التحرر من السرعة. حيث حددت العامل الأول للسرعة: هو عدم وصول المفحوصين إلى (15% أو 10% أو 5%) من مجموع فقرات الاختبار. والعامل الثاني للسرعة: هو أسلوب الإجابة عن الفقرات التي لم يتم الوصول إليها (بأن يترك المفحوصون الفقرات فارغة، أو أن يجيبوا عنها بشكل عشوائي، أو خليط من نمطي الإجابة ترك الفقرة فارغة/ الإجابة العشوائية). والعامل الثالث من عوامل السرعة: ترتيب الفقرات في الاختبار (ترتيبا عشوائيا أو ترتيبها من الأسهل إلى الأصعب)، وتم دراسة آثار هذه العوامل على تقدير معلمة القدرة، ومعالم الفقرات، حيث أظهرت النتائج أن تقدير القدرة كان الأقل تأثرا في انتهاك افتراض التحرر من السرعة، إلا أن التأثير كان جوهرياً على تقديرات معالم الفقرة، حيث أعطت تقديرات متحيزة للأعلى لمعاملات الصعوبة (أي أصعب) والتمييز (أكثر تمييزا) (Overestimation)، وتقديرات متدنية لمعلمة التخمين (Underestimation)، خصوصا لدى الفقرات الواقعة في نهاية الاختبار(كونها محذوفة حكما أو بشكل تلقائي).

وبطريقة مشابهة لطريقة اوشيما (Oshima, 1994) هدفت دراسة سجينبك (Schnipke,1999) إلى استقصاء درجة تشويه السرعة لتقديرات معالم الصعوبة باستخدام

بيانات زمن الاستجابة، وذلك عن طريق توليد بيانات الاستجابة الصحيحة والخاطئة وزمن الاستجابة مع وجود تخمين سريع (بمعنى آخر، عندما يقارب زمن الاختبار على الانتهاء فإن بعض المفحوصين يخمنون الإجابة عن بقية الفقرات المتبقية ظنا منهم أن يحصلوا على بعض الإجابات الصحيحة منها بالحظ وذلك ما يسمى بالتخمين السريع) أو مع عدم وجود تخمين سريع للمفحوصين، وبعد ذلك تم مقارنة تقديرات معالم نظرية استجابة الفقرة ومنحنيات خصائص الفقرة مع بعضها البعض وثم مقارنتها مع قيمها الحقيقية. ووجدت الباحثة بأنه عندما يكون التخمين السريع موجودا في عملية تقدير المعالم، فإن الفقرات التي تقع في نهاية الاختبار تبدو أكثر صعوبة، وأن الدقة تكون دالة بمكان الفقرة؛ بمعنى آخر إذا كان المفحوصون قد استجابوا عن الفقرات من بداية الاختبار حتى نهايته بشكل متسلسل، فإن استجاباتهم تكون اقل دقة كلما اقتربوا من نهاية الاختبار. وأيضا عندما لا يوجد تخمين سريع فانه يمكن أن تقدر معلمة صعوبة الفقرة (b) بشكل أفضل ودقيق. ولا يمكن تقدير معلمة التمييز (a) بشكل جيد كتقدير معلمة الصعوبة بوجود تخمين سريع. ولا يمكن أن تقدر معلمة التخمين (c) بشكل أفضل كتقدير معلمة التمييز بوجود التخمين السريع.

وجدت الدراستان السابقتان: دراسة اوشيما (Oshima, 1994) ودراسة سجينبك (Schnipke,1999) الآثار التي ترجع إلى السرعة ولكن الدراسة الأخيرة دراسة جينبك (Schnipke,1999) أضافت جزء سلوك التخمين السريع واستخدمت بيانات زمن الاستجابة.

ثانيا: الدراسات التي بحثت في أثر زمن الاختبار على أداء المفحوصين

هدفت دراسة لواين (Wayne, 2003) إلى مراجعة عدة دراسات تناولت تأثير اختلاف زمن اختبار المواضيع الكتابية في اختبار (SAT II) على علامات الطلبة بمختلف مستويات قدراتهم وتلك الدراسات هي الآتية:

قام ليفينغستون (Livingston, 1987) بدراسة هدفت إلى دراسة الفروق في علامات الطلبة في المواضيع الكتابية، التي تم إكمالها في ثلاثة أوقات زمنية مختلفة: إما إكمال الموضوع بزمن 20 دقيقة، أو إكمال الموضوع بزمن 30 دقيقة، أو إكمال الموضوع بزمن (10) دقائق منفصلة لوضع خطة لكتابة الموضوع، و20 دقيقة لكتابة الموضوع. ولتحقيق هدف الدراسة تم استخدام موضوعين مختلفين.

وبينت النتائج التي وضعت حول زمن الموضوع أن أثر إضافة (10) دقائق (أي السماح بوقت (30) دقيقة بدلا من (20) دقيقة لكتابة الموضوع) "كان ضئيلا جدا". وأجري تحليل إضافي لمعرفة اثر تمديد زمن الكتابة على مستويات قدرة الطلبة: بالنسبة للطلبة ذوي القدرة المتدنية، لم يكن للوقت الإضافي، ولا للموضوع أي اختلاف في علاماتهم، أما بالنسبة للطلبة ذوي القدرة المتوسطة ؛ أي الذين تقع بين علاماتهم (6) و(8) على تدريج العلامات (2– 12)، فكان لتمديد الوقت (10) دقائق أن ارتفعت علاماتهم من (10/1) إلى (6/1) جزء من نقطة في تدريج العلامات (2–12). وبالنسبة للطلبة ذوي القدرة العالية، أظهرت العشرة دقائق الإضافية تحسنا في العلامات بمعدل 2/1 نقطة.

وفي دراسة قام بها كرون ورايت وبارون (Crone, Wright & Baron ,1993) هدفت إلى التحقق من تأثير طول الموضوع الكتابي لتحديد الزمن النهائي لكتابة المواضيع في اختبار (SAT II) الكتابي. ولتحقيق ذلك اجري اختبار على (7100) من طلبة المدارس الثانوية أكملوا موضوعين كتابيين بزمن (30) دقيقة آو بزمن (15) دقيقة. وبينت النتائج بوضوح أن الطلبة قد أحرزوا علامات اقل عند إعطائهم الموضوع بزمن (15) دقيقة من العلامات التي حصلوا عليها في الموضوع الذي كان زمنه (30) دقيقة. وكان الاختلاف حوالي (1.22) نقطة في مقياس العلامات ((2−12)). وخلص الباحثون إلى القول بأن الطلبة قادرون على كتابة موضوع مناسب أجنبية (TOEFL)ولكن بجودة أقل مما لو كتبوه في 30 دقيقة، وأن هنالك تأثيرا قليلا على الدرجات الرتبية للطلبة.

وأجرت هيل (Hale, 1992) دراسة في اختبار اللغة الإنجليزية كلغة أجنبية (TOEFL) هدفت إلى مقارنة تأثير إطالة زمن الاختبار على أداء المفحوصين، ولتحقيق ذلك الهدف أعطي (820) مفحوصا من الناطقين بغير اللغة الإنجليزية مواضيع كتابية بزمن (30) دقيقة، أو(45) دقيقة. وقد أشارت نتائج الدراسة إلى أن معامل الارتباط بين علامات المفحوصين الذين أتموا كتابة المواضيع بزمن (30) دقيقة وعلامات المفحوصين الذين أتموا كتابة المواضيع بزمن (45) دقيقة كان (0.77). وزاد الوقت الإضافي العلامات بمقدار (3/1) انحراف معياري، وكذلك كان له أثرا قليل في ترتيب علامات المفحوصين، ولكن لا يوجد اختلاف ذو دلالة في آثاره في أداء المفحوصين ذوي القدرة الضعيفة والقدرة العالية.

وفي دراسة لباورز وفاولز (Powers & Fowles, 1996) هدفت إلى فحص اثر اختلاف الزمن المحدد في الجزء الكتابي من اختبار (GRE) على أداء المفحوصين . وقد تكونت عينة الدراسة من (300) مفحوص من طلبة الدراسات العليا شاركوا في كتابة موضوعين مختلفين ضمن زمنين مختلفين الأول: الوقت المحدد المعياري (40) دقيقة والثاني: (60) دقيقة؛ أي تمديد الوقت المعياري. وجد الباحثان بأنه يمكن تحسين أداء المفحوصين على الاختبار عندما يعطوا (60) دقيقة بدلا من (40) دقيقة لإكمال الاختبار، حيث كان متوسط الزيادة في علاماتهم ما بين (0.06) و(1.0) علامة على تدريج العلامات (1-6) من قبل قارئين اثنين.

وخلص واين (Wayne, 2003) إلى مراجعة الدراسات السابقة الذكر من إلى القول بأن الدراسات التي تم مراجعتها كان محورها ثبات وصدق وصعوبة اختبارات المواضيع الكتابية المعطاة ضمن ظروف زمنية مختلقة، أشارت إلى أن إعطاء المفحوصين مزيدا من الوقت لإكمال كتابة الموضوع ربما يرفع من علاماتهم ، ولكن لا يغير في ترتيب الطلبة، ولا توجد مؤشرات واضحة تدل على أن إعطاء الطلبة وقتا أقل لإكمال الموضوع يحسن، أو ينتقص من أي مستوى من مستويات قدرة المفحوصين .

وأجرى مانديناج وبريجمان و كاهلان ليتيوسيس و ترباني (Mandinach, Bridgeman, Cahalan-Laitusis & Trapani, 2005) دراسة كان هدفها بيان أثر تمديد زمن الاختبار في أداء المفحوصين في اختبار (SAT). وذلك عن طريق إعطاء المفحوصين وقتا معياريا، وإعطائهم زمنا ونصف الزمن المعياري مع استراحات أو دونها، أو ضعف

الوقت المعياري دون جزء مخصص للاستراحة بين جزئي الاختبار. ويمكن تلخيص نتائج هذه الدراسة كالآتي: استفاد المفحوصون ذوو القدرة المتدنية قليلا من الوقت الإضافي، ومن المعروف أن أي زمن يمكن وضعه لا يحسن من أداء المفحوصين إذا كانوا لا يملكون المعرفة أو المهارة. وساعدت الاستراحات في الاختبار المفحوصين بكافة مستوياتهم المختلفة من القدرة. واستفاد المفحوصون ذوو القدرة المتوسطة والعالية من تمديد الوقت.

وفي دراسة لبريجمان وكلاين وهيسنجر (Bridgeman, Cline & Hessinger 2003) التي هدفت إلى التحقق من أن سرعة الاستجابة تلعب دورا مهما في اختبارات (GRE) الاختبار العام (General Test)، وذلك بإضافة جزء من الاختبار في قسم المصطلحات أو في القسم الكمي من اختبار (GRE)، إما بنفس الوقت المعياري المحدد، أو بمرة ونصف المرة من الوقت المعياري المحدد. دلت النتائج على أن العلامات التي تم الحصول عليها من (15948) مفحوصا أشارت إلى أن الوقت الإضافي رفع حوالي (7) علامات (في مقياس العلامات (200-800)) لعلامات المفحوصين فـي جزء المصطلحات و الجزء الكمي، وكان المفحوصين ذوي القدرة المتدنية الأكثر تأثرا بتمديد زمن الاختبار، حيث كانت علامتهم أعلى قليلا عند تمديد زمن الاختبار.

التعقيب على الدراسات السابقة

إن الاهتمام بفرض التحرر من السرعة من قبل المشتغلين في القياس التربوي بشكل عام، والقياس النفسي والعقلي بشكل خاص يدل على أهمية الافتراض عند تطوير المقاييس والاختبارات في معظم المجالات، وتؤكد على ذلك الدراسات التي أجريت في هذا المجال.

وفي ضوء ذلك يتضح من هذه الدراسات تعدد طرق تقدير سرعة الاختبار، فقد كان هنالك طرق تقليدية لتقدير السرعة إما طرق تستخدم تطبيق واحد للاختبار مثل نسبة جلكسن وطريقة مؤسسة الاختبارات التربوية، أو تطبيق متعدد للاختبار مثل طريقة كرونباخ وورينغتون، كما تم استخدام في تقدير السرعة إما بيانات حقيقية للاختبار مثل اختبار SAT أو اختبار GRE أو اختبار التوفل TOEFL أو استخدمت بيانات محاكاة تم توليد فيها بيانات اختبارات مشابهة لبيانات الاختبارات الحقيقية.

بالإضافة إلى ذلك وجد بأن معظم الدراسات التي بحثت في اثر انتهاك افتراض السرعة على تقديرات معالم الأفراد والفقرة ركزت على دراسة عدة عوامل لانتهاك الافتراض منها: نسبة الفقرات التي لم يكمل المفحوصون الإجابة عليها أو ترتيب الفقرات من الأسهل إلى الأصعب أو كيفية الإجابة على الفقرات التي تقع قبيل انتهاء الاختبار أو أن يخمنوا في الإجابة أو أن يتركوها فارغة، وكما تم استخدام زمن الاستجابة في تقدير المعالم. واستخدمت في هذه الدراسات برامج حاسوبية متقدمة لتقدير المعالم، إلا أنها أغفلت الدراسات عامل الزمن وأثره على معالم الفقرة والأفراد وهذا ما ستناقشه الدراسة الحالية حيث اعتمدت على عامل واحد لانتهاك افتراض التحرر من السرعة لم تتطرق إليها الدراسات السابقة بشيء من التفصيل ألا وهو زمن الاختبار، ودراسة أثر ذلك العامل على تقديرات معالم الأفراد والفقرات واستخدم في هذه الدراسة برنامجا حاسوبيا متقدما من اجل تقدير تلك المعالم .

وفيما يخص دراسات اثر تمديد زمن الاختبار على أداء المفحوصين فقد اقتصرت الدراسات على بيان أثر الزمن على أداء المفحوصين في المواضيع الكتابية وذلك لقلة الدراسات

التي تناولت أثر الزمن على أداء المفحوصين في اختبارات الاختيار من متعدد (Armitage,1999)، أما الدراسة الحالية استخدمت الاختبار المتعدد الاختيارات.

كما أن هنالك دراسات عديدة ناقشت اثر تمديد زمن الاختبار في أداء المفحوصين كما ورد سابقا حيث لم يستفيد المفحوصون كثيرا من إطالة زمن الاختبار، فقد رفع من علاماتهم ولكن لم يؤثر ذلك الارتفاع في ترتيب علامات المفحوصين؛ أي لم يؤثر طول الزمن على مدلول علامات المفحوصين، إلا أنه لا توجد دراسات ناقشت اثر تقليل زمن الاختبار على أداء المفحوصين و,(Wayne 2003) وتقديرات معالم الفقرات ولأفراد، وتم مناقشة هذا الجانب في الدراسة الحالية.

الفصل الثالث

الطريقة والإجراءات

يتناول هذا الفصل وصفا لمجتمع الدراسة وطريقة اختيار العينة، وبناء أداة الدراسة، والنماذج الإحصائية المستخدمة، والتحقق من افتراضات النموذجين من النماذج اللوجستية لنظرية استجابة الفقرة، والنماذج الإحصائية المستخدمة، والإجراءات التي اتبعت في مطابقة البيانات مع النماذج الإحصائية المستخدمة، وكذلك وصفا للمعالجات الإحصائية التي تم إجراءها من أجل الحصول على إجابات أسئلة الدراسة.

مجتمع الدراسة

تكوَّن مجتمع الدراسة من جميع طلبة الصف التاسع الأساسي من الذكور والإناث والمنتظمين في المدارس الحكومية والخاصة التابعة لكل من مديرية التربية والتعليم لمنطقة عمان الرابعة، وعمان الخامسة للعام الدراسي 2012/2011، الموزعين على (29) مدرسة في مديرية تربية عمان الرابعة و(44) مدرسة في مديرية تربية عمان الخامسة.

عينة الدراسة

تم اختيار عينة الدراسة بالطريقة العشوائية، حيث اشتملت عينة الدراسة على (1500) طالب وطالبة من طلبة الصف التاسع الأساسي في (18) مدرسة من مديريتي التربية والتعليم لمنطقة عمان الرابعة ومنطقة عمان الخامسة. وتم تطبيق الاختبار على جميع الشعب الموجودة في كل مدرسة من المدارس التي تم اختيارها؛ لأن علامة الاختبار تم احتسابها ضمن علامات

الطلبة المدرسية، وذلك لضمان جدية الطلبة في الإجابة على فقرات الاختبار. وكان عدد الطلبة الذين شملهم الاختبار 1502 طالب وطالبة، مما استدعى حذف ورقتي إجابة بطريقة عشوائية، كانتا من نصيب الصورتين: الصورة الثانية للاختبار (إعطاء الطلبة 60 دقيقة لإكمال الاختبار)، والصورة الثالثة للاختبار (إعطاء الطلبة 75 دقيقة لإكمال الاختبار)، وذلك للحصول على أعداد متساوية من الطلبة لكل صورة من صور الاختبار الثلاث، حيث أصبح عدد أفراد العينة التي حللت استجاباتهم (500) طالب وطالبة لكل صورة من صور الاختبار الثلاث.

أداة الدراسة

من أجل تحقيق أهداف الدراسة، تم بناء اختبار تحصيلي في وحدة الهندسة الإحداثية ووحدة النسب المثلثية في مبحث الرياضيات للصف التاسع الأساسي،وكانت فقرات هذا الاختبار جميعها من نوع الاختيار من متعدد، بأربعة بدائل لكل فقرة. وقد اقتضى بناء هذا الاختبار ووضعه في صورته النهائية القيام ببعض الإجراءات التي يمكن إجمالها بالخطوات الآتية:

أولا: تحديد الغرض من الاختبار، وهو قياس تحصيل الطلبة في وحدة الهندسة الإحداثية ووحدة النسب المثلثية في مادة الرياضيات للصف التاسع الأساسي بافتراض أن المعارف أو المهارات في هاتين الوحدتين تقيسان سمة أو قدرة أحادية البعد (Unidimensional).

ثانيا: كتابة أهداف المحتوى، وعمل جدول المواصفات الذي تم فيه ربط مستويات الأهداف بمحتوى المادة الدراسية موضوع الاختبار. وتوضح الملاحق (1)،(2)،(3)،(4) جدول المواصفات الذي بني عليه الاختبار إضافة إلى توزيع عدد الأسئلة.

ثالثا: كُتبت (40) فقرة من نوع الاختيار من متعدد لكل فقرة منها أربعة بدائل أحدها صحيح، وقد روعيت الأسس الفنية في كتابة هذا النوع من الفقرات، من حيث ملاءمتها للهدف الذي تقيسه في ضوء المحتوى والمستوى المعرفي، كما روعي أن لا تعتمد إجابة إحدى هذه الفقرات على إجابة الفقرات الأخرى.

رابعا: تم عرض الاختبار بصورته الأولية (40 فقرة) على مجموعة من المحكمين من أصحاب الاختصاص والخبرة؛ للتأكد من صدق المحتوى للاختبار، وطلب من كل منهم على حدة أن يبدي رأيه حول دقة الصياغة، ومدى تمثيل الفقرات للأهداف، وبناء على أرائهم تم إجراء التعديلات اللازمة على فقرات الاختبار، وقد تم حذف الفقرات التي أجمع المحكمون على أن هناك فقرات أخرى تؤدي نفس الغرض، وبناء على ذلك تم حذف (5) فقرات، وتعديل الفقرات المتبقية وصياغتها بشكل دقيق، حيث بلغ عدد فقرات الاختبار حتى هذه المرحلة (35) فقرة.

خامسا: التجريب الأولي للاختبار على عينة استطلاعية

تم تطبيق الاختبار بصورته الأولية على مجموعة مكونة من (50) طالب وطالبة من إحدى المدارس الخاصة التي تم اختيارها عشوائيا من مديرية تربية منطقة عمان الخامسة تم اختيارهما بشكل عشوائي من شعب الصف التاسع الأساسي في المدرسة. وبالاعتماد على

التجريب الأولي للاختبار تم تقدير الزمن المناسب للإجابة عن فقرات الاختبار بحـــوالي (60) دقيقة للإجابة على فقرات الاختبار، وذلك بعد الحصول على مؤشرات سيرد ذكرها فـــي بنـــد افتراض التحرر من السرعة (Speededness) صفحة (36). ثم أجريت عملية التحليل لبيانات التجريب الأولي باستخدام برنامج التحليل الإحصائي (SPSS)، حيث تم حساب درجة صعوبة الفقرة بإيجاد نسبة عدد الطلبة الذين أجابوا عن الفقرة إجابة صحيحة من بين الطلبـــة الـــذين حاولوا الإجابة عن هذه الفقرة. وتم احتساب القدرة التمييزية للفقرة من خلال حساب معاملات الارتباط بوينت بايسيريال بين الدرجات على الفقرة والدرجة الكلية على الاختبار، وكان ذلـــك بالاعتماد على مفاهيم النظرية التقليدية في القياس، وذلك لأن عدد أفراد العينة الاســـتطلاعية (50) طالبا وطالبة، وهو غير كاف لاستخدام نموذج أحادي المعلمة (نموذج راش) أو نموذج ثنائي المعلمة، حيث إن استخدام نموذج أحادي المعلمة (نموذج راش) يتطلـــب حجـــم عينـــة (200) فرد كحد أدنى، كما أشار بيجار (Bejar, 1980)، واستخدام نموذج ثنـــائي المعلمـــة يتطلب حجم عينة (500) فرد كحد أدنى. وقد تم حذف (5) فقرات اثنتان منهما لأن معامل صعوبتها أقل من (0.2)، و(3) فقرات لأن معامل الارتباط بينها وبين الاختبار أقل من (0.2) وفق ما أشار إليه كل من ألن وين (Allen & Yen, 1979) وكروكر والجينا (Crocker & Algina, 1986)، وبذلك أصبح لدى الباحث اختبار بصورته النهائية مكون من (30) فقـــرة جاهزة للتطبيق على العينة كما هو واضح في ملحق(4) .

سادسا: تم إنقاص طول زمن الاختبار المناسب (60 دقيقة) بمقدار ربع الساعة ليكون زمن الاختبار (45 دقيقة) (الصورة الأولى للاختبار)، وزيادة طول زمن الاختبار المناســـب

المخصص للاختبار بمقدار ربع الساعة ليكون طول الاختبار (75 دقيقة) (الصورة الثالثة للاختبار)، وأما الصورة الثانية، فكانت من نصيب طول الوقت المناسب الذي تم اختياره (60 دقيقة).

سابعا: إعداد ورقة تعليمات للاختبار وورقة إجابة نموذجية لكل صورة من صور الاختبار الثلاث.

ثامنا: إجراءات جمع البيانات:

تم تطبيق الاختبار في الفترة الزمنية الممتدة بين 2012/3/8 -2012/5/8،(وتشمل هذه الفترة منذ موافقة وزارة التربية والتعليم على إجراء تطبيق الاختبار وتطبيق الاختبار وتصحيحه) وذلك كما يأتي:

- تم الاستعانة بمدرسي مادة الرياضيات للصف التاسع الأساسي في كل مدرسة من مديريات التربية التي شملتها العينة، حيث تم اطلاعهم على الاختبار وفقراته. وقد تم تسليم كل مدرس مغلفا يتكون من كراسة الاختبار ونموذج الإجابة، وتم التأكد من أن جميع كراسات الاختبار كاملة بجميع صفحاتها، وأنها مرقمة ومصورة بطريقة جيدة ومقروءة، وزود المدرسون برقم هاتف الباحث؛ للاستفسار عن أية أمور طارئة في أثناء عملية التطبيق، وكما شارك الباحث في إدارة الاختبار في بعض المدارس.
- بعد تطبيق صور الاختبار الثلاث بصورتها النهائية على عينة الدراسة، تم جمع استجابات الطلبة عن صور الاختبار، وتمت مراجعة أوراق الطلبة؛ للكشف عن الاستجابات ذات النمطية الواحدة، ليتم استثناؤها من التحليل، وبناء على ذلك لم تستثنَ أي ورقة من التحليل.

النماذج الإحصائية المستخدمة

1- النموذج اللوجستي أحادي المعلمة (نموذج راش) (One – Parameter Logistic Model).

2- النموذج اللوجستي ثنائي المعالم (Two - Parameter Logistic Model).

التحقق من افتراضات النموذجين من النماذج اللوجستية لنظرية اســتجابة الفقــرة،

النموذج اللوجستي أحادي المعلمة (نموذج راش) (One – Parameter Logistic Model) والنموذج اللوجستي ثنائي المعالم (Two – Parameter Logistic Model).

- فحص أحادية البعد للبيانات (unidimensionality): هنالك عدد من المؤشرات يمكن أن تستخدم كأدلة على أحادية البعد وهذه المؤشرات هي: التحليل العاملي، ونمط الاستجابة، والاتساق الداخلي، والمكونات الرئيسية.

ويمكن التحقق من تلك المؤشرات على النحو الآتي:

1– التحليل العاملي:

اقترح لورد (Lord,1980) دلالة إحصائية تقريبية في حساب الجذور الكامنة في التحليل العاملي لتحقيق أحادية البعد لمجموعة فقرات الاختبار،وهي كون قيمة الجذر الأول كبيرة مقارنة بالجذر الثاني والجذر الثاني قريبا في قيمته من الجذور الأخرى، الذي يدل بدوره على أن الفقرات أحادية البعد.

2– نمط الاستجابة:

ويتم ذلك من خلال ترتيب الفقرات وفق صعوبتها، فالمفحوص الذي يفشل في الإجابة عن فقرة بمستوى صعوبة معين يفشل في الإجابة عن فقرة بمستوى أعلى.

3- الاتساق الداخلي:

يُعد معامل ارتباط الفقرة بالدرجة الكلية مؤشرا على أن الفقرات تقيس سمة واحدة. ويشير نانلي (Nannally) في هذا الصدد إلى أن أهم إحصائي في بناء الاختبار هو معامل ارتباط الفقرة بالدرجة الكلية. فالفقرات ذات أعلى معامل ارتباط (بوينت بايسيريال) تستبقى في الاختبار، لأن مثل هذه الفقرات عند تجميعها تؤلف تدريج على اتساق داخلي، وهذا يعطي مؤشرا إلى أن مجموعة الفقرات تقيس سمة أو قدرة واحدة.

4- المكونات الرئيسية:

يتم بناء اختبار بافتراض أحادية البعد أولا، ثم إجراء التحليل العاملي بطريقة المكونات الرئيسية وحذف الفقرات التي لا تقيس العامل العام (البخاتي، 2009).

وتم في الدراسة الحالية استخدام الطريقة الأكثر شيوعا في الدراسات طريقة التحليل العاملي.

تم التحقق من هذا الافتراض (افتراض أحادية البعد) باستخدام برنامج (SPSS) لكل صورة من صور الاختبار الثلاث، حيث تم إجراء أسلوب التحليل العاملي (Factor Analysis) باستخدام تحليل المكونات الرئيسية (Principal Component Analysis)

لاستجابات الأفراد عن صور الاختبار الثلاث، وقد جرى التدوير باستخدام طريقة التدوير المتعامد (Varimax Rotation)للعوامل التي كانت قيم الجذر الكامن لها أكبر من (1)، وقد تم حساب الجذور الكامنة (Eigenvalues) ونسبة التباين المفسر (Explained Variance) لكل عامل من العوامل، ونسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني وملاحظة نسبة الفرق بين الجذر الكامن الأول والثاني إلى الفرق بين الجذر الكامن الثاني والثالث، وذلك لصور الاختبار الثلاث. ويتضح كمؤشر ثاني لتحقق افتراض أحادية البعد من خلال التمثيل البياني باستخدام (Scree Plot) لقيم الجذور الكامنة لكل صورة من صور الاختبار، ووجود عامل سائد على بقية العوامل.

- فحص افتراض الاستقلال الموضعي (Local Independence)، وهـذا الافتـراض مكافيء لافتراض أحادية البعد، كما أوصى به هامبلتون وسوامينثان (Hamleton & Swaminathan, 1985).

- فحص التحرر من السرعة (Speededness)، تم تقدير الزمن الأولي للاختبـار بعـد التجريب الأولي له، حيث حدد الباحث زمن الاختبار بساعة بعـد أن حـصل علـى المؤشرات الآتية:

أ- بلغت نسبة الطلبة الذين أكملوا جميع فقرات الاختبار (85%).

ب- بلغت نسبة الطلبة الذين أنهوا (75%) من فقرات الاختبار (99%).

ج-بلغ عدد الفقرات التي أكملها (80%) من الطلاب (23) فقرة.

واستخدمت في هذه الدراسة كثيرا من المعالجات الإحصائية، التي يمكن تصنيفها كما يأتي:

المعالجات الإحصائية لمطابقة البيانات مع النموذج الأحادي المعلمة (نموذج راش).

باستخدام برنامج (BILOG-MG) تم ما يأتي:

- استخدام هذا البرنامج في تحليل البيانات وفق نموذج راش، إن هذا البرنامج يعطي تحليلا يتركز حول قدرة المفحوص، بحيث يعطي نتائج على شكل النتائج التي يعطيها نموذج ثلاثي المعلمة، حيث تكون قيم معاملات التمييز ثابتة لجميع الفقرات، وقيم معامل التخمين صفر لها أيضا.
- يستخدم الإحصائي كاي تربيع (χ^2) لاختبار مطابقة الفقرة للنموذج، حيث يتم تقسيم الدرجات إلى (20) فئة على الأكثر ويوضع في كل فئة خمسة مفحوصين على الأقل، لذا يتوقع أن يستخدم عددا للفئات أقل من (20)، والعدد نفسه لهذه الفئات سوف يستخدم لجميع الفقرات بحيث تكون درجات الحرية تساوي (عدد الفئات – 1).
- إن قيما عالية للإحصائي كاي تربيع (χ^2) تعكس عدم مطابقة الفقرة مع النموذج المستخدم. وعموما إذا تجاوزت قيم الإحصائي كاي تربيع (χ^2) القيمة الحرجة له، فإن الضعف في مطابقة الفقرة للنموذج سيكون عند مستوى الدلالة (0.05).
- وبالتالي فإن قيما منخفضة للدلالة الإحصائية تقل عن مستوى الدلالة (α=0.05) لكل فقرة تعكس عدم مطابقة الفقرة للنموذج.

المعالجات الإحصائية لمطابقة البيانات مع النموذج اللوجستي ثنائي المعالم (Two - Parameter Logistic Model).

استخرجت معاملات صعوبة وتمييز فقرات الاختبار بصوره الثلاث من برنامج (BILOG - MG)، حيث يستخدم هذا البرنامج الإحصائي كاي تربيع (χ^2) لاختبار مطابقة الفقرة للنموذج.

- أما الفقرات التي يقع معامل صعوبتها خارج المدى من−1.5 إلى 1.5 فقد اعتبرت ضعيفة، والفقرات التي تتمتع بقدرة تمييزية خارج المدى من0.4 إلى 2، اعتبرت ضعيفة، أو أن قيما منخفضة للدلالة الإحصائية تقل عن مستوى الدلالة (α=0.05) لكل فقرة تعكس عدم مطابقة الفقرة للنموذج.

الكشف عن الفروق في تقدير معالم الفقرات والأفراد لصور الاختبار الثلاث

للكشف عن الفروق في دقة تقدير معلمة صعوبة الفقرات لصور الاختبار الثلاث باستخدام نموذج راش والنموذج لثنائي المعلمة، تم إجراء تحليل التباين الأحادي لمتوسطات الأخطاء المعيارية في تقدير هذه المعالم باستخدام برنامج (SPSS). واعتمادا على وجود أو عدم وجود فروق ذات دلالة إحصائية بين متوسطات الأخطاء المعيارية في تقدير صعوبة الفقرات تم الحكم على أي الصور من صور الاختبار أكثر دقة في تقدير صعوبة الفقرات، وذلك باستخدام طريقة شيفيه (Scheffe's Method) لعقد المقارنات البعدية بين متوسطات الأخطاء المعيارية لتقديرات صعوبة الفقرات.

وللكشف عن الفروق في دقة تقدير تمييز الفقرات لصور الاختبار الثلاث باستخدام النموذج الثنائي المعالم، تم إجراء تحليل التباين الأحادي لمتوسطات الأخطاء المعيارية في تقدير هذه المعالم باستخدام برنامج (SPSS). واعتمادا على وجود أو عدم وجود فروق ذات دلالة إحصائية بين متوسطات الأخطاء المعيارية في تقدير تمييز الفقرات تم الحكم على أي الصور من صور الاختبار أكثر دقة في تقدير تمييز الفقرات.

وللكشف عن الفروق في دقة تقدير قدرات الأفراد لصور الاختبار الثلاث باستخدام نموذج راش والنموذج لثنائي المعلمة، تم إجراء تحليل التباين الأحادي لمتوسطات الأخطاء المعيارية في تقدير هذه المعالم باستخدام برنامج (SPSS). واعتمادا على وجود أو عدم وجود فروق ذات دلالة إحصائية بين متوسطات الأخطاء المعيارية في تقدير قدرات الأفراد تم الحكم على أي الصور من صور الاختبار أكثر دقة في تقدير قدرة الأفراد وذلك باستخدام طريقة عقد المقارنات البعدية بين متوسطات الأخطاء المعيارية لتقديرات قدرات الافراد.

وتم إيجاد دالة المعلومات للفقرات والاختبار لصور الاختبار الثلاث باستخدام نموذج راش والنموذج لثنائي المعلمة باستخدام برنامج (BILOG-MG) المصمم لتحليل البيانات وفق نماذج استجابة الفقرة، كما تم الحكم على إعطاء أي صورة من صور الاختبارات الثلاث أعلى قدر من المعلومات عند مستويات القدرة لدى الأفراد ذوي القدرات المنخفضة، أو ذوي القدرات المتوسطة، أو ذوي القدرات العالية.

الفصل الرابع

نتائج الدراسة ومناقشتها

يقدم هذا الفصل وصفا للنتائج التي توصل إليها الباحث، ومناقشة هذه النتائج، وما يرافق ذلك من استنتاجات وتوصيات تسهم في إثراء البحث. ويمكن عرض نتائج الدراسة وفقا للترتيب الآتي:

1- النتائج المتعلقة بالتحليل العاملي، وذلك للتأكد من تحقق فرض أحادية البعد للاختبار.

2- النتائج المتعلقة بمطابقة الاستجابات عن فقرات كل صورة من صور الاختبار الثلاث.

3- النتائج المتعلقة بدقة تقدير معلمة صعوبة الفقرة باستخدام النموذج الأحادي المعلمة (نموذج راش) لصور الاختبار الثلاث.

4- النتائج المتعلقة بدقة تقدير معلمة صعوبة الفقرة باستخدام النموذج الثنائي المعلمة لصور الاختبار الثلاث.

5- النتائج المتعلقة بدقة تقدير معلمة تمييز الفقرة باستخدام النموذج الثنائي المعلمة لصور الاختبار الثلاث.

6- النتائج المتعلقة بدقة تقدير معلمة قدرة الأفراد باستخدام النموذج الأحادي المعلمة (نموذج راش) لصور الاختبار الثلاث.

7- النتائج المتعلقة بدقة تقدير معلمة قدرة الأفراد باستخدام النموذج الثنائي المعلمة لصور الاختبار الثلاث.

8- النتائج المتعلقة بقيم دالة المعلومات للفقرة والاختبار باستخدام النموذج الأحادي المعلمة (نموذج راش) لصور الاختبار الثلاث.

9- النتائج المتعلقة بقيم دالة المعلومات للفقرة والاختبار باستخدام النموذج الثنـــائي المعلمـــة لصور الاختبار الثلاث.

أولا: النتائج المتعلقة بالتحليل العاملي للتأكد من افتراض أحادية البعد للاختبار

تم إجراء تحليل عاملي للاستجابات (Factor Analysis) باستخدام المكونات الأساســية (Principal Component)، حيث حسبت قيمة الجذر الكامن(EigenValue)، ونــسبة التبــاين المفسر (Explained Variance) لكل عامــل مــن العوامــل، إذ أشــار هــامبلتون وتــروب (Hambleton & Traub, 1971) إلى إمكانية استخدام هذا النوع من التحليل لفحــص أحاديــة البعد.

نتائج التحليل العاملي لصورة الاختبار الأولى (طول زمن الاختبار (45) دقيقة)

أوضحت نتائج التحليل العاملي لاستجابات الصورة الأولى للاختبار أن قيمة الجذر الكامن للعامل الأول قد بلغت (4.57)، وفسر من التباين ما نسبته (15.25%)، في حين كانت قيمة الجذر الكامن الثاني (2.37)، وفسر من التباين ما نسبته (7.90%) فقط.

و يبين الجدول (1) قيم الجذور الكامنة، ونسب التباين المفسر لأول إحدى عشر عاملا، وكــذلك نسب التباين المفسر التراكمية.

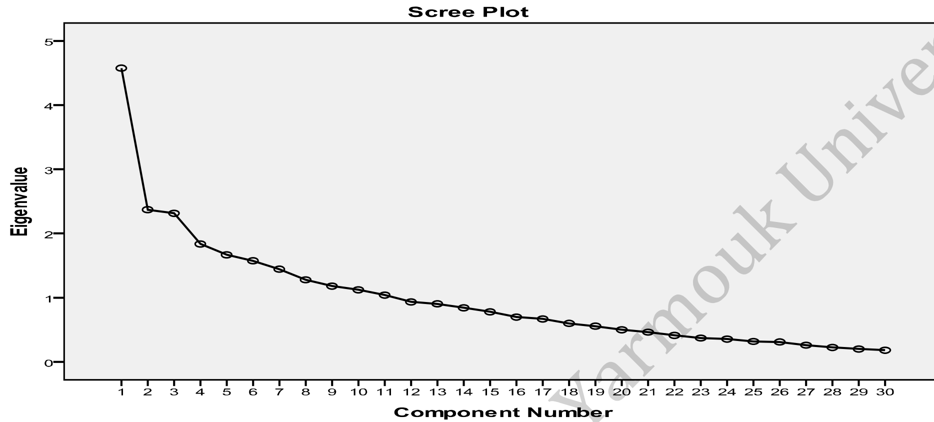
الجدول (1)

قيم الجذور الكامنة ونسب التباين المفسر ونسب التباين المفسر التراكمي لصورة الاختبار الأولى (إعطاء الطلاب 45 دقيقة للإجابة عن فقرات الاختبار)

العامل	الجذر الكامن	التباين المفسر	التباين المفسر التراكمي
1	4.57	15.25	15.25
2	2.37	7.90	23.15
3	2.32	7.72	30.87
4	1.84	6.12	36.99
5	1.67	5.56	42.55
6	1.58	5.25	47.80
7	1.44	4.81	52.61
8	1.28	4.26	56.87
9	1.18	3.94	60.81
10	1.12	3.74	64.55
11	1.04	3.47	68.02

ويتضح من الجدول (1) أن هنالك (3) عوامل فقط كانت قيم الجذر الكامن لها أكبر من (2)، وفسرا معا ما نسبته (30.87%) من التباين، كما يشير الجدول إلى أن مجموع ما فسره أول أحد عشر عاملا كان حوالي (68.02%). وكمؤشر على أحادية البعد، يتبين أن قيمة الجذر الكامن الأول كانت أقل من ضعف قيمة الجذر الكامن الثاني بقليل، وبالاعتماد على ما أشار إليه ريكس (Reckase , 1979) وما ذكره لورد(lord) (المشار إليه في دراسة البنيز وفورسيث،Albanese& Forsyth 1984) يظهر أن نسبة الجذر الكامن الأول إلى الجذر الكامن الثاني تساوي (1.928=2.37/4.57) وهي اقل من (2)، والذي يعتبر مؤشرا على عدم تحقق افتراض أحادية البعد (Hambleton & Swaminithan, 1985; Alastair & Hutchinson, 1987). وهذا قد يعني ظهور عامل آخر هو عامل السرعة.

كما يوضح الشكل (1) التمثيل البياني باستخدام (Scree Plot) لقيم الجذور الكامنة للعوامل المكونة لصورة الاختبار الأولى.



الشكل (1): التمثيل البياني لقيم الجذور الكامنة للعوامل المكونة لصورة الاختبار الأولى (إعطاء الطلاب 45 دقيقة للإجابة عن فقرات الاختبار)

نتائج التحليل العاملي لصورة الاختبار الثانية (طول زمن الاختبار (60) دقيقة)

أوضحت نتائج التحليل العاملي لاستجابات الصورة الثانية للاختبار أن قيمة الجذر الكامن للعامل الأول قد بلغت (6.15)، وفسر من التباين ما نسبته (20.51%) في حين كانت قيمة الجذر الكامن الثاني (2.43)، وفسر من التباين ما نسبته (8.09%) فقط. والجدول (2) يبين قيم الجذور الكامنة، ونسب التباين المفسر لأول عشرة عوامل، وكذلك نسب التباين المفسر التراكمية.

الجدول (2)

قيم الجذور الكامنة ونسب التباين المفسر ونسب التباين المفسر التراكمي لصورة الاختبار الثانية (إعطاء الطلاب 60 دقيقة للإجابة عن فقرات الاختبار)

	الجذر الكامن	التباين المفسر	التباين المفسر التراكمي
1	6.15	20.51	20.51
2	2.43	8.09	28.60
3	2.25	7.50	36.10
4	2.03	6.77	42.87
5	1.68	5.60	48.47
6	1.66	5.54	54.01
7	1.45	4.82	58.83
8	1.30	4.34	63.17
9	1.11	3.69	66.86
10	1.03	3.44	70.30

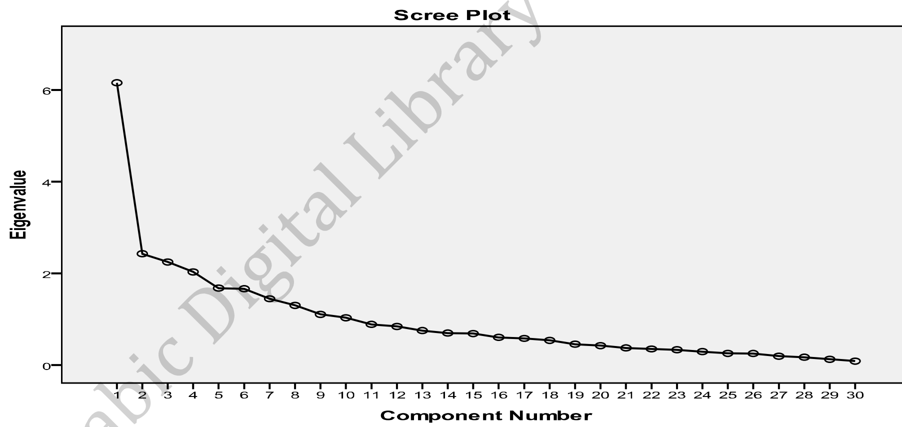
ويتضح من الجدول (2) أن هنالك أربعة عوامل فقط كانت قيم الجذر الكامن لها أكبر من (2)، وفسرا معا ما نسبته (42.87%) من التباين، ويشير الجدول إلى أن مجموع ما فسرته أول عشرة عوامل كان حوالي (70.30%).

وكمؤشر على أحادية البعد يظهر أن قيمة الجذر الكامن الأول، قد كانت حوالي ضعفين ونصف الضعف، أو أقل بقليل من قيمة الجذر الكامن الثاني؛ أي أن نسبة الجذر الكامن الأول إلى الجذر الكامن الثاني تساوي (2.53=2.43/6.15)، وهي أكبر من (2)، الذي اعتبر مؤشرا لأحادية البعد.

كما ويمكن الإشارة إلى أن جميع قيم الجذور الكامنة للعوامل الأخرى من الخامس إلى العاشر، قد كانت متقاربة وصغيرة جدا، وتراوحت بين (1.03– 1.68)، وهذا دليل آخر على

أحادية البعد؛ لأنها تمثل مجموعة متماثلة ومتقاربة من العوامل الصغيرة باستثناء العامل الأول الذي تفوق بشكل ملحوظ.

كما يتضح تحقق افتراض أحادية البعد من خلال التمثيل البياني باستخدام (Scree Plot) لقيم الجذور الكامنة للعوامل المكونة لصورة الاختبار الثانية الموضح في الشكل (2)، والذي يبين أن قيمة الجذر الكامن للعامل الأول تميز بشكل واضح عن الجذور الكامنة لبقية العوامل، وذلك لأن نسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني كانت عالية، وهذا مؤشر آخر على تحقق افتراض أحادية البعد.



الشكل (2): التمثيل البياني لقيم الجذور الكامنة للعوامل المكونة لصورة الاختبار الثانية (إعطاء الطلاب 60 دقيقة للإجابة عن فقرات الاختبار)

نتائج التحليل العاملي لصورة الاختبار الثالثة (طول زمن الاختبار (75) دقيقة)

أوضحت نتائج التحليل العاملي لاستجابات الصورة الثالثة للاختبار أن قيمة الجذر الكامن للعامل الأول قد بلغت (7.07)، وفسر من التباين ما نسبته (23.56%) في حين كانت قيمة الجذر الكامن الثاني (2.24) وفسر من التباين ما نسبته (7.46%) فقط. والجدول (3) يبين قيم الجــذر الكامن ونسب التباين المفسر لأول تسعة عوامل، وكذلك نسب التباين المفسر التراكمية.

الجدول (3)

قيم الجذور الكامنة ونسب التباين المفسر ونسب التباين المفسر التراكمي لصورة الاختبار الثالثة (إعطاء الطلاب 75 دقيقة للإجابة عن فقرات الاختبار)

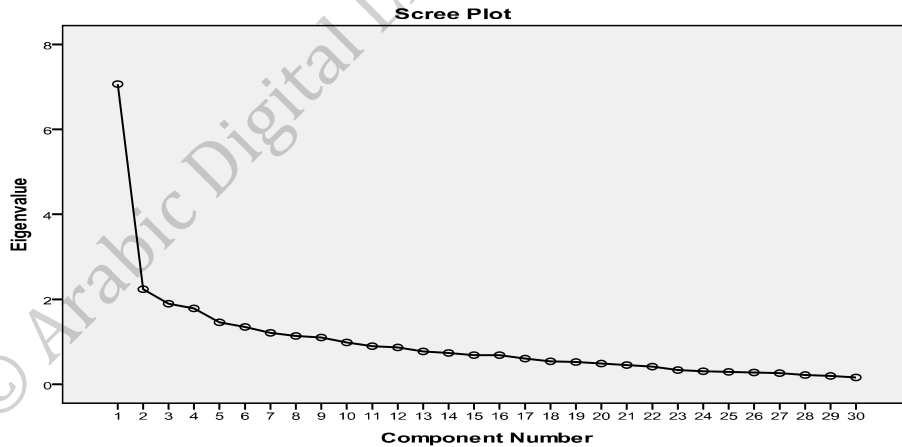
العامل	الجذر الكامن	التباين المفسر	التباين المفسر التراكمي
1	7.07	23.56	23.56
2	2.24	7.46	31.02
3	1.90	6.33	37.36
4	1.79	5.96	43.31
5	1.46	4.85	48.17
6	1.35	4.51	52.68
7	1.21	4.05	56.72
8	1.14	3.80	60.52
9	1.11	3.68	64.20

يتضح من الجدول أن هنالك عاملين فقط كانت قيم الجذر الكامن لها أكبر من (2)، وفسرا معا ما نسبته (31.02%) من التباين، ويشير الجدول إلى أن مجموع ما فسرته أول تسعة عوامل كان حوالي (64.20%). وكمؤشر على أحادية البعد يظهر أن قيمة الجذر الكــامن الأول كانــت حوالي ثلاث أضعاف قيمة الجذر الكامن الثاني؛ أي أن نسبة الجذر الكــامن الأول إلــى الجــذر

الكامن الثاني تساوي (3.16=2.24/7.07)، وهي أكبر من (2)، واعتبر هذا مؤشـــرا لأحاديـــة البعد.

كما يمكن الإشارة إلى أن جميع قيم الجذور الكامنة للعوامل الأخرى مـــن الثالـــث إلـــى التاسع قد كانت متقاربة وصغيرة جدا، و تراوحت بين (1.11–1.90) وهذا دليل آخر على أحادية البعد؛ لأنها تمثل مجموعة متماثلة ومتقاربة من العوامل الصغيرة، باستثناء العامل الأول الذي كان عاليا.

كما يتضح تحقق افتراض أحادية البعد من خلال التمثيل البياني باســتخدام (Scree Plot) لقيم الجذور الكامنة للعوامل المكونة لصورة الاختبار الثالثة الموضح في الشكل (3)، والذي يبين أن العامل الأول كان سائدا على بقية العوامل، وهذا مؤشر آخر على تحقق افتراض أحادية البعد.



الشكل (3): التمثيل البياني لقيم الجذور الكامنة للعوامل المكونة لصورة الاختبار الثالثة (إعطاء الطلاب 75 دقيقة للإجابة عن فقرات الاختبار)

ملخص نتائج التحايل العاملي لصور الاختبار الثلاث

يبين الجدول (4) ملخصا لنتائج التحليل العاملي لصور الاختبار الثلاث ، حيث يوضح الجدول قيم الجذور الكامنة ونسب التباين المفسر لأول عاملين حسب صور الاختبار، وناتج قسمة الجذر للعامل الأول على العامل الثاني لكل صورة من صور الاختبار.

جدول (4): ملخص نتائج التحليل العاملي لصور الاختبار الثلاث

الصورة	الجذر الكامن الأول	التباين المفسر	الجذر الكامن الثاني	التباين المفسر	نسبة الجذر الكامن الأول إلى الجذر الكامن الثاني
زمن 45 دقيقة	4.57	15.25	2.37	7.90	1.93
زمن 60 دقيقة	6.15	20.51	2.43	8.09	2.53
زمن 75 دقيقة	7.07	23.56	2.24	7.46	3.16

من الجدول (4) أوضحت نتائج التحليل العاملي لاستجابات صور الاختبار الثلاث أن قيم الجذور الكامنة للعامل الأول وللعامل الثاني أكبر من (2)، وهذا مؤشر على تحقق افتراض أحادية البعد. إلا أن نسبة الجذر الكامن الأول إلى الجذر الكامن الثاني في الصورة الأولى للاختبار كانت تساوي (1.93) وهي اقل من (2)، والذي يعتبر مؤشرا على عدم تحقق افتراض أحادية البعد، وهذا قد يعني ظهور عامل آخر هو عامل السرعة.

ثانيا: النتائج المتعلقة بمطابقة الاستجابات عن كل صورة من صور الاختبار

معلمة الصعوبة المقدرة باستخدام النموذج اللوجستي الأحادي المعلمة (نموذج راش) حسب الصورة الأولى للاختبار

من أجل مطابقة فقرات الاختبار بصورته الأولى مع النموذج اللوجستي أحادي المعلمة (نموذج راش) تم تحليل بيانات استجابات الطلبة عن فقرات الإختبار باستخدام برنامج حاسوبي (BILOG-MG)، حيث تم من خلاله تقدير معلمة الصعوبة، وحساب الخطأ المعياري لتقدير معلم الصعوبة، وحساب قيمة الإحصائي كاي تربيع χ^2 وحساب الدلالة الإحصائية في حسن المطابقة، والجدول (5) يوضحها.

جدول (5) : معالم الصعوبة المقدرة باستخدام النموذج اللوجستي الأحادي للصورة الأولى للاختبار

الدلالة الإحصائية	χ^2	الخطأ المعياري لتقدير معلم الصعوبة	معلم الصعوبة	رقم الفقرة
.000	37.3	.146	.144	1
.027	17.3	.140	-.093	2
.004	21.1	.145	.237	3
.084	15.3	.134	.383	4
.550	7.8	.132	.012	5
.528	8.1	.136	-.106	6
.422	9.2	.138	.329	7
.125	12.6	.138	.210	8
.001	27.6	.146	.436	9
.000	32.5	.128	.409	10
.000	33.6	.145	.091	11
.003	23.2	.146	.517	12
.000	44.1	.147	.197	13
.795	5.4	.135	-.040	14
.717	6.2	.135	-.093	15
.108	13.1	.140	.104	16
.009	20.5	.145	.383	17
.657	6.8	.134	.396	18
.104	14.5	.131	.223	19
.180	12.6	.133	-.001	20
.000	53.0	.149	-.119	21
.121	14.0	.133	.223	22
.260	11.2	.137	.557	23
.925	3.8	.136	-.054	24
.000	47.9	.148	.157	25
.004	24.2	.130	.250	26
.579	7.6	.135	.557	27
.000	44.3	.148	-.014	28
.586	6.6	.139	.197	29
.304	9.5	.141	.436	30

يتضح من الجدول (5) أن قيم معلمة الصعوبة تتراوح من (−0.119) إلى (0.557) مما يعني أن القدرات ضعيفة. ومتوسطها يساوي (0.191) والانحراف المعياري لها يساوي (0.200).

وبملاحظة قيم الخطأ المعياري لتقدير معلمة الصعوبة وجد الباحث أن أعلى قيمة له كانت (0.149) عند الفقرة (21)، وهي الفقرة السهلة جدا. وحيث أن قيم الدلالة الإحصائية تعبر عن مطابقة الفقرة للنموذج؛ أي أن أي قيمة الدلالة الإحصائية عند كل فقرة تقل عن مستوى الدلالة الاحصائية (0.05) تعكس عدم مطابقة الفقرة للنموذج، وبذا تكون الفقرات (1،2،3،9،10،11،12،13،17،21،25،26،28) غير مطابقة مع النموذج لأن قيم الدلالة الإحصائية لها تقل عن (0.05) ؛ أي أن عدد الفقرات الغير مطابقة وفق نموذج راش (13) فقرة أي ما نسبته (43%).

معالم الصعوبة المقدرة باستخدام النموذج اللوجستي الثنائي المعالم للصورة الثانية للاختبار

تم تحليل البيانات باستخدام برنامج (BILOG-MG)، وتم التركيز على معلمة الصعوبة، والجدول (6) يوضح معالم الصعوبة المقدرة والخطأ المعياري في تقدير معالم الصعوبة وقيم الإحصائي كاي تربيع χ^2 وحساب الدلالة الإحصائية في حسن المطابقة.

جدول(6): معالم الصعوبة المقدرة باستخدام النموذج اللوجستي الثنائي للصورة الأولى للاختبار

رقم الفقرة	معلم الصعوبة	الخطأ المعياري لتقدير معلم الصعوبة	χ^2	الدلالة الإحصائية
1	.05	.033	7.2	.028
2	-.15	.195	52.5	.000
3	.14	.093	173.0	.000
4	.56	.235	10.5	.309
5	.01	.267	7.3	.606
6	-.23	.294	25.0	.003
7	.49	.237	31.0	.000
8	.29	.213	70.0	.000
9	.29	.104	201.0	.000
10	.93	.390	10.1	.339
11	.02	.034	2.1	.553
12	.34	.105	173.8	.000
13	.07	.033	6.9	.031
14	-.09	.246	23.5	.005
15	-.23	.338	18.0	.035
16	.13	.200	60.3	.000
17	.26	.105	207.4	.000
18	.73	.310	10.6	.304
19	.63	.434	34.1	.000
20	-.02	.271	22.5	.007
21	-.06	.034	2.2	.335
22	.40	.278	32.4	.000
23	.81	.260	14.5	.106
24	-.12	.286	24.8	.003
25	.06	.032	7.4	.025
26	.54	.336	26.9	.001
27	.97	.319	28.1	.001
28	-.02	.033	3.9	.139
29	.28	.214	62.3	.000
30	.33	.115	185.5	.000

ومن الجدول (6) وجد أن قيم معلمة الصعوبة تتراوح من (−0.23) إلى (0.97) وقيمة متوسط معلمة الصعوبة تساوي (0.247)، والانحراف المعياري لها يساوي (0.332).

وبملاحظة قيم الخطأ المعياري في الجدول وجد الباحث أن أعلى قيمة له كانت (0.434)، والتي تقابل الفقرة (19) السهلة جدا. ولأن أي قيمة للدلالة الإحصائية تقل عن الدلالة الإحصائية (0.05) تعكس عدم مطابقة الفقرة للنموذج، وبناء عليه تكون الفقرات (1،2،3،6،7،8،9،12،13،14،15،16،17،19،20،22،24،25،26،27،29،30) غير مطابقة مع النموذج؛ أي أن عدد الفقرات غير المطابقة وفق النموذج الثنائي (22) فقرة؛ أي ما نسبته (70%).

معالم الصعوبة المقدرة باستخدام النموذج اللوجستي أحادي المعلمة للصورة الأولى للاختبار

اظهر تحليل البيانات عند مطابقة الفقرات مع النموذج اللوجستي أحادي المعلمة للصورة الثانية للاختبار قيم معالم الصعوبة وقيم الخطأ المعياري لتقدير معلم الصعوبة وقيم الإحصائي كاي تربيع χ^2 والدلالة الإحصائية لتقدير معالم الصعوبة والجدول (7) يوضحها.

جدول (7): معالم الصعوبة المقدرة باستخدام النموذج اللوجستي الأحادي للصورة الثانية للاختبار

رقم الفقرة	معلم الصعوبة	الخطأ المعياري لتقدير معلم الصعوبة	χ^2	الدلالة الإحصائية
1	-1.88	.261	8.8	.455
2	-.92	.252	10.3	.329
3	-1.80	.262	21.2	.007
4	-1.61	.258	12.5	.132
5	-1.58	.260	2.9	.890
6	-1.27	.250	20.2	.010
7	-.05	.251	16.2	.063
8	-1.53	.261	25.9	.002
9	-1.34	.259	16.5	.035
10	-.89	.252	4.7	.857
11	-.66	.247	19.9	.011
12	.80	.246	42.4	.000
13	-2.24	.272	19.5	.012
14	-1.75	.268	21.4	.003
15	-2.40	.274	9.2	.325
16	-1.73	.263	6.1	.636
17	-1.39	.260	7.8	.346
18	-.23	.253	8.4	.399
19	-1.91	.266	8.3	.501
20	-2.90	.289	29.1	.000
21	-1.58	.261	24.8	.003
22	-1.48	.261	9.2	.323
23	-.66	.252	7.8	.453
24	-1.66	.267	42.5	.000
25	-2.29	.275	26.8	.002
26	.27	.242	33.5	.000
27	-1.15	.256	8.6	.378
28	-1.73	.264	9.5	.219
29	-1.39	.260	18.6	.017
30	-1.22	.259	21.5	.006

كانت قيم معلمة الصعوبة تتراوح من (−2.9) إلى (0.8) وكان متوسط معلمة الصعوبة تساوي (−1.339) والانحراف المعياري لها يساوي (0.782).

وبملاحظة قيم الخطأ المعياري في الجدول (7) وجد أن أعلى قيمة له كانــت (0.289)، وهي الفقرة (20) الأكثر سهولة.

ومـــــــــن الجـــــــــدول (7) وجـــــــــد أن الفقـــــــــرات (26،25،24،21،20،14،13،12،11،9،8،6،3،29،30) غير مطابقة مع النموذج لأن قيم الدلالة الإحصائية لها تقل عن (0.05)؛ أي أن عدد الفقرات غير المطابقة وفق نموذج راش (15) فقرة؛ أي ما نسبته (50%).

معالم الصعوبة المقدرة باستخدام النموذج اللوجستي ثنائي المعالم للصورة الثانية للاختبار

ولمطابقة الفقرات مع النموذج اللوجستي ثنائي المعالم للصورة الثانية للاختبار، تم تحليل البيانات كما هو موضح في الجدول (8).

جدول (8): معالم الصعوبة المقدرة باستخدام النموذج اللوجستي الثنائي للصورة الثانية للاختبار

رقم الفقرة	معلم الصعوبة	الخطأ المعياري لتقدير معلم الصعوبة	χ^2	الدلالة الإحصائية
1	-1.91	.507	16.2	.064
2	-.91	.313	5.5	.786
3	-1.51	.372	11.7	.112
4	-1.58	.425	16.0	.067
5	-1.20	.295	5.6	.695
6	-2.11	.688	17.6	.040
7	-.04	.166	34.6	.000
8	-1.07	.253	14.3	.075
9	-.77	.172	12.2	.144
10	-.90	.317	4.2	.895
11	-1.02	.467	36.8	.000
12	1.43	.576	51.6	.000
13	-1.52	.319	4.1	.904
14	-.71	.116	14.2	.028
15	-1.53	.308	1.2	.992
16	-1.23	.281	6.1	.531
17	-.90	.218	6.4	.498
18	-.14	.139	7.0	.541
19	-1.25	.267	3.3	.911
20	-1.12	.149	23.0	.001
21	-1.11	.258	11.8	.163
22	-.87	.192	2.1	.956
23	-.60	.252	9.5	.298
24	-.72	.134	7.1	.310
25	-1.07	.169	29.1	.000
*26				
27	-.98	.288	8.1	.423
28	-.98	.205	6.8	.455
29	-.87	.206	6.0	.540
30	-.77	.197	13.8	.088

*تم استبعاد الفقرة 26 من التحليل لان معامل تمييزها صغير جدا وبالتالي تم استبعادها من قبل البرنامج الإحصائي المستخدم(BILOG –MG).

ومن الجدول (8) يتبين أن قيم معلمة الصعوبة تتراوح من (−2,11) إلى (1.43) وأن متوسط معلمة الصعوبة تساوي (−0.964) والانحراف المعياري لها يساوي (0.628).

ومقابل كل فقرة أظهرت النتائج قيمة الخطأ المعياري في تقدير معلمة الصعوبة، وبملاحظة قيمه في الجدول وجد أن أعلى قيمة له كانت (0.688)، وهي الفقرة (6) السهلة جدا.

وأظهرت الجدول (8) أيضا قيم الإحصائي كاي تربيع χ^2 وقيم الدلالة الإحصائية. هذا ويكون عدد الفقرات الغير مطابقة وفق النموذج اللوجستي ثنائي المعالم (8) فقرات، هي الفقرات (26،25،20،14،12،11،5،6) أي ما نسبته (24%)، (مع الأخذ بعين الاعتبار الفقرة المحذوفة رقم (26))

معالم الصعوبة المقدرة باستخدام النموذج اللوجستي الأحادي المعلمة للصورة الثالثة للاختبار

تم تحليل البيانات باستخدام برنامج (BILOG-MG)، وتم التركيز على معلمة الصعوبة، وكانت قيم معلمة الصعوبة تتراوح من (−2.71) إلى (−0.58) مما يعني أن الفقرات سهلة جدا، والجدول (9) يبين قيم الصعوبة للفقرات المقدرة حسب النموذج اللوجستي أحادي المعلمة (نموذج راش)، وكان متوسط معلمة الصعوبة تساوي (−1.753) والانحراف المعياري لها يساوي (0.461).

جدول (9): معالم الصعوبة المقدرة باستخدام النموذج اللوجستي أحادي المعلمة

(نموذج راش) للصورة الثالثة للاختبار

رقم الفقرة	معلم الصعوبة	الخطأ المعياري لتقدير معلم الصعوبة	χ^2	الدلالة الإحصائية
1	-1.77	.208	.004	.004
2	-1.97	.206	.828	.828
3	-2.09	.207	.341	.341
4	-1.48	.199	.915	.915
5	-2.01	.208	.556	.556
6	-2.45	.222	.231	.231
7	-1.36	.199	.077	.077
8	-1.69	.201	.635	.635
9	-1.36	.200	.022	.022
10	-.58	.179	.000	.000
11	-1.85	.206	.003	.003
12	-1.40	.195	.383	.383
13	-1.48	.196	.780	.780
14	-2.01	.209	.133	.133
15	-2.66	.224	.801	.801
16	-1.91	.209	.142	.142
17	-1.34	.195	.508	.508
18	-1.77	.201	.937	.937
19	-1.87	.207	.264	.264
20	-2.14	.211	.514	.514
21	-1.67	.198	.510	.510
22	-1.75	.203	.032	.032
23	-1.29	.193	.678	.678
24	-1.51	.200	.072	.072
25	-2.33	.221	.000	.000
26	-1.09	.190	.279	.279
27	-1.73	.201	.951	.951
28	-2.71	.225	.812	.812
29	-2.01	.208	.171	.171
30	-1.55	.197	.103	.103

تم تحليل البيانات باستخدام برنامج (BILOG-MG)، وتم التركيز على معلمة الصعوبة، وكانت قيم معلمة الصعوبة تتراوح من (−2.71) إلى (−0.58)، الجدول (9) يبين أن متوسط معلمة الصعوبة يساوي (−1.753) والانحراف المعياري لها يساوي (0.461).

ومقابل كل فقرة أظهرت النتائج قيمة الخطأ المعياري فـــي تقـــدير معلمـــة الـــصعوبة، وبملاحظة قيمه في الجدول وجد أن أعلى قيمة له كانت (0.225)، وهي عند الفقرة (28) السهلة جدا. وأظهرت النتائج أيضا قيم الإحصائي كاي تربيع χ^2، و الدلالة الإحصائية. ومن الجدول (9) وجد أن الفقرات (25،22،10،9،8،1) غير مطابقة مع النموذج لأن قيم الدلالة الإحصائية لها تقل عن (0.05) ، أي أن عدد الفقرات غير المطابقة وفق نموذج راش (6) فقرات؛ أي مـــا نـــسبته (%20).

معالم الصعوبة المقدرة باستخدام النموذج اللوجستي ثنائي المعالم للصورة الثالثة للاختبار

تم تحليل البيانات باستخدام برنامج (BILOG-MG)، وتم التركيز على معلمة الصعوبة، وكانت قيم معلمة الصعوبة تتراوح من (−2.31) إلى (−0.95). كما يوضحه الجدول (10).

جدول (10): معالم الصعوبة باستخدام النموذج اللوجستي الثنائي للصورة الثالثة للاختبار

رقم الفقرة	معلم الصعوبة	الخطأ المعياري لتقدير معلم الصعوبة	χ^2	الدلالة الإحصائية
1	-.95	.137	18.7	.009
2	-1.97	.425	7.5	.383
3	-2.31	.524	12.0	.150
4	-1.08	.208	4.7	.697
5	-1.55	.278	3.0	.888
6	-1.43	.210	7.3	.294
7	-.94	.181	6.3	.500
8	-1.46	.301	4.8	.782
9	-.89	.168	4.3	.829
10	-1.45	.610	24.2	.002
11	-1.13	.166	27.7	.000
12	-1.62	.396	28.8	.001
13	-1.62	.385	10.9	.207
14	-1.51	.275	8.9	.350
15	-2.08	.374	5.6	.687
16	-1.29	.221	3.4	.847
17	-1.35	.315	12.2	.144
18	-1.44	.272	11.4	.183
19	-1.45	.276	3.5	.834
20	-1.74	.330	3.4	.843
21	-2.18	.539	10.9	.206
22	-1.37	.259	9.9	.196
23	-1.37	.330	3.1	.929
24	-1.10	.207	5.3	.723
25	-1.14	.134	19.1	.004
26	-1.55	.445	21.9	.009
27	-1.86	.418	9.7	.284
28	-2.19	.409	4.7	.702
29	-1.66	.319	12.2	.144
30	-1.50	.322	2.5	.962

وكان متوسط معلمة الصعوبة تساوي (−1.526) والانحراف المعياري لها يساوي (0.361) ومقابل كل فقرة أظهرت النتائج قيمة الخطأ المعياري في تقدير معلمة الصعوبة، وبملاحظة قيمه في الجدول (10) وجد أن أعلى قيمة له كانت (0.610)، عند الفقرة (10) السهلة جدا.

وأظهرت نتائج التحليل أيضا قيم الإحصائي كاي تربيع χ^2، وقيم الدلالـــة الإحـــصائية. ومن الجدول (10) وجد أن الفقرات (1،10،11،12،25،26) غير مطابقة مع النموذج لأن قـــيم مستوى الدلالة الإحصائية لها يقل عن (0.05) ؛ أي أن عدد الفقرات الغير مطابقة وفق النموذج الثنائي المعالم (6) فقرات؛ أي ما نسبته (20%).

معالم التمييز المقدرة باستخدام النموذج اللوجستي ثنائي المعلمة للصورة الأولى للاختبار

تم تحليل البيانات باستخدام برنامج (BILOG-MG)، وتم التركيز على معلمة التمييـــز، كما هو موضح في جدول (11).

جدول (11): معالم التمييز المقدرة باستخدام النموذج اللوجستي الثنائي للصورة الأولى للاختبار

رقم الفقرة	معلم التمييز	الخطأ المعياري لتقدير معلم التمييز
1	4.27	.518
2	.29	.055
3	.66	.110
4	.26	.049
5	.20	.043
6	.19	.042
7	.25	.051
8	.27	.054
9	.64	.106
10	.16	.038
11	3.45	.359
12	.66	.107
13	3.96	.495
14	.22	.046
15	.16	.037
16	.27	.055
17	.60	.099
18	.20	.044
19	.13	.031
20	.20	.043
21	4.56	.495
22	.21	.044
23	.26	.052
24	.19	.043
25	4.22	.520
26	.17	.039
27	.22	.046
28	4.11	.453
29	.26	.053
30	.55	.089

ومن الجدول (11) يتبين أن قيم معلم التمييز تتراوح بين (0.13) و(4.56)، حيث كـان أعلى تمييز للفقرة (21)، وأن الفقرات (28،25،21،13،11،1) غير مطابقة لان معامل تمييزها يقع خارج الفترة المقبولة [0،2]، كما أشار (علام، 1986) " إنه نادرا ما تم الحصول عمليا على قيم لمعامل التمييز أكبر من (2) وحدة معيارية، لذلك فإن المدى المعتاد لقيم معامل التمييز يقع في الفترة [0،2]". وكان متوسط معلم التمييز يساوي (1.060) والانحراف المعيـاري لـه يـساوي (1.533).

معالم التمييز المقدرة باستخدام النموذج اللوجستي ثنائي المعلمة للصورة الثانية للاختبار

تم تحليل البيانات باستخدام برنامج (BILOG-MG)، وتم التركيز على معلـم التمييـز، الجدول (12) يوضح ذلك.

جدول (12): معالم التمييز المقدرة باستخدام النموذج اللوجستي الثنائي عند الصورة الثانية للاختبار

الفقرة	معلم التمييز	الخطأ المعياري لتقدير معلم التمييز
1	.21	.051
2	.22	.050
3	.27	.060
4	.22	.052
5	.30	.063
6	.13	.032
7	.34	.060
8	.33	.067
9	.42	.074
10	.22	.049
11	.14	.034
12	.12	.030
13	.34	.067
14	.68	.109
15	.37	.074
16	.32	.068
17	.36	.071
18	.41	.073
19	.35	.068
20	.71	.126
21	.33	.062
22	.41	.077
23	.24	.054
24	.61	.103
25	.54	.084
27	.26	.057
28	.42	.079
29	.37	.074
30	.37	.072

كانت قيم معلمة التمييز للفقرات تتراوح من (0.12) إلى (71. 0))، حيث كان أعلى تمييز للفقرة (20)، وعليه، فإن معظم الفقرات تقع ضمن المدى المقبول [2،0]، وكان متوسط التمييز يساوي (0.345)، والانحراف المعياري يساوي (0.145).وكانت تمييز الفقرة (26) اقل من الفترة المقبولة حيث كانت قيمة التمييز لها (0.18-) هذا وقد تم حذفها من قبل البرنامج (BILOG– MG).

معالم التمييز المقدرة باستخدام النموذج اللوجستي ثنائي المعلمة للصورة الثالثة للاختبار

تم تحليل البيانات باستخدام برنامج (BILOG-MG)، وتم التركيز على معلم التمييز والجدول (13) يوضح ذلك.

جدول (13): معالم التمييز المقدرة باستخدام النموذج اللوجستي الثنائي عند الصورة الثالثة للاختبار

رقم الفقرة	معلم التمييز	الخطأ المعياري لتقدير معلم التمييز
1	.64	.096
2	.30	.061
3	.26	.058
4	.43	.076
5	.40	.072
6	.57	.096
7	.46	.076
8	.35	.067
9	.50	.080
10	.11	.027
11	.54	.080
12	.25	.051
13	.27	.055
14	.41	.074
15	.39	.073
16	.47	.082
17	.29	.060
18	.37	.066
19	.40	.073
20	.38	.072
21	.22	.050
22	.39	.074
23	.28	.057
24	.43	.075
25	.75	.113
26	.20	.045
27	.27	.056
28	.38	.074
29	.37	.071
30	.31	.059

كانت قيم معلمة التمييز للفقرات تتراوح من (0.11) إلى (0.75)، حيث كان أعلى تمييز للفقرة (25). وعليه، فإن معظم الفقرات تقع ضمن المدى المقبول [2،0] وكان متوسط التمييــز يساوي (0.380)، والانحراف المعياري يساوي (0.132).

ثالثا: النتائج المتعلقة بدقة تقدير معالم الفقرات والأفراد لصور الاختبار الثلاث

النتائج المتعلقة بالسؤال الفرعي الأول (هل تختلف دقة تقدير معلمة صعوبة الفقرة باســتخدام النموذج الأحادي المعلمة (نموذج راش) لصور الاختبار الثلاث باختلاف زمن الاختبار؟)

تم حساب المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم الصعوبة لفقرات كــل صورة من صور الاختبار.كما هو واضح في جدول (14).

جدول (14):المتوسطات الحسابية والانحرافات المعيارية للخطأ لمعياري لتقدير معلم الصعوبة في النموذج الأحادي وفقا لمتغير وقت الاختبار

الوقت	العدد	المتوسط الحسابي للخطأ المعياري	الانحراف المعياري
45	30	0.139	0.006
60	30	0.260	0.010
75	30	0.204	0.010
الكلي	90	0.201	0.050

ويتبين من الجدول (14) أن المتوسطات الحسابية للأخطاء المعيارية لتقــديرات معــالم الصعوبة لفقرات كل صورة من صور الاختبار (0.139، 0.260،0.204) حسب ترتيب صور الاختبار.

ولفحص دقة تقديرات معالم الصعوبة، فقد تم استخدام تحليل التباين الأحادي للأخطاء المعيارية لهذه التقديرات لكل من صور الاختبار الثلاث، والجدول (15) يوضح ذلك.

جدول (15): نتائج تحليل التباين الأحادي للخطأ لمعياري لتقدير معلم الصعوبة في النموذج الأحادي وفقا لمتغير وقت الاختبار

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	ف	الدلالة الإحصائية
وقت الاختبار	.220	2	.110	1406.746	.001
الخطأ	.007	87	.000		
الكلي	.227	89			

ويلاحظ من الجدول (15) وجود فروق ذات دلالة إحصائية في تقدير صعوبة الفقرات تعزى إلى صور الاختبار، حيث كان متوسط الأخطاء المعيارية للصورة الثانية أعلى منه للصورتين الأولى والثالثة، ولذلك تم استخدام اختبار شيفيه (Scheffe's Method) للكشف عن مواضع الفرق بين متوسطات الأخطاء المعيارية لتقدير الصعوبة لفقرات الصور الثلاث للاختبار لمعرفة أي الفرق دالا إحصائيا. ويبين الجدول (16) نتائج المقارنات الثنائية بين متوسطات الأخطاء المعيارية.

جدول (16): المقارنات البعدية بين متوسطات الخطأ المعياري لتقدير معلم الصعوبة في النموذج الأحادي وفقا لمتغير وقت الاختبار

المقارنات	الفرق بين المتوسطات	Std. Error	الدلالة الإحصائية
45 * 60	-0.121	0.002	.001
45* 75	-0.065	0.002	.001
60 * 75	0.056	0.002	.001

ويلاحظ من الجدول (16) أن الفرق بين متوسطي الأخطاء المعيارية كان دالا إحصائيا بين الصورتين الأولى والثانية ولصالح الصورة الثانية؛ أي أن فقرات الصورة الأولى أكثر دقة في تقديرات الصعوبة من فقرات الصورة الثانية. كما أشارت إلى أن الفرق بين متوسطي الأخطاء المعيارية كان دالا إحصائيا بين الصورتين الأولى والثالثة ولصالح الصورة الثالثة؛ أي أن فقرات الصورة الأولى للاختبار أكثر دقة في تقديرات معالم الصعوبة من فقرات الصورة الثالثة للاختبار. كما أشارت النتائج إلى أن الفرق بين متوسطي الأخطاء المعيارية كان دالا إحصائيا بين الصورتين الثانية والثالثة ولصالح الصورة الثانية ؛ أي أن فقرات الصورة الثالثة للاختبار أكثر دقة في تقديرات معالم الصعوبة من فقرات الصورة الثانية للاختبار.

- **النتائج المتعلقة بالسؤال الفرعي الأول (هل تختلف دقة تقدير معلمة صعوبة الفقرة باستخدام النموذج الثنائي المعلمة لصور الاختبار الثلاث باختلاف زمن الاختبار؟)**

تم حساب المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم الصعوبة لفقرات كل صورة من صور الاختبار، والموضحة في الجدول (17).

جدول (17):المتوسطات الحسابية والانحرافات المعيارية للخطأ لمعياري لتقدير معلم الصعوبة في النموذج الثنائي وفقا لمتغير وقت الاختبار

الوقت	العدد	المتوسط الحسابي للخطأ المعياري	الانحراف المعياري
45	30	0.201	0.119
60	29	0.284	0.137
75	30	0.314	0.121
الكلي	89	0.267	0.133

كانت المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم الصعوبة لفقرات كل صورة (0.314,0.284,0.201) حسب الترتيب.

ولفحص دقة تقدير معالم الصعوبة، فقد تم استخدام تحليل التباين الأحادي للأخطاء المعيارية لهذه التقديرات لكل من صور الاختبار الثلاث، والموضحة في الجدول (18).

جدول (18): نتائج تحليل التباين الأحادي للخطأ لمعياري لتقدير معلم الصعوبة في النموذج الثنائي وفقا لمتغير وقت الاختبار

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	ف	الدلالة الإحصائية
وقت الاختبار	.205	2	.103	6.488	.002
الخطأ	1.360	86	.016		
الكلي	1.565	88			

ويلاحظ من الجدول (18) وجود فرق ذا دلالة إحصائية بين تقدير صعوبة الفقرات يعزى إلى صور الاختبار، حيث كان متوسط الأخطاء المعيارية للصورة الثالثة أعلى منه للصورة الأولى والثانية، ولذلك تم استخدام اختبار المقارنات البعدية بين متوسطات الأخطاء المعيارية لتقدير الصعوبة لفقرات الصور الثلاث للاختبار لمعرفة أي الفرق دالا إحصائيا. ويبين الجدول (19) نتائج المقارنات الثنائية بين متوسطات الأخطاء المعيارية.

جدول (19): المقارنات البعدية بين متوسطات الخطأ المعياري لتقدير معلم الصعوبة في النموذج الثنائي وفقا لمتغير وقت الاختبار

المقارنات	الفرق بين المتوسطات	Std. Error	الدلالة الإحصائية
45 * 60	-0.083	0.033	.001
45* 75	-0.113	0.032	.001
60 * 75	-0.030	0.033	.001

ويلاحظ من الجدول (19) أن الفرق بين متوسطي الأخطاء المعيارية كان دالا إحصائيا بين الصورتين الأولى، والثالثة، ولصالح الصورة الثالثة؛ أي أن فقرات الصورة الأولى للاختبار أكثر دقة في تقديرات معالم الصعوبة من فقرات الصورة الثالثة للاختبار. كما أشارت النتائج إلى أن الفرق بين متوسطي الأخطاء المعيارية كان دالا إحصائيا بين الصورتين الأولى والثانية ولصالح الصورة الثانية ؛ أي أن فقرات الصورة الأولى أكثر دقة في تقديرات الصعوبة من فقرات الصورة الثانية. كما أشارت النتائج إلى أنه يوجد دلالة إحصائية للفرق بين متوسطي الأخطاء المعيارية بين الصورتين الثانية والثالثة لصالح الصورة الثالثة؛ أي أن فقرات الصورة الثانية للاختبار أكثر دقة في تقديرات معالم الصعوبة من فقرات الصورة الثالثة للاختبار.

رابعا: النتائج المتعلقة بالسؤال الفرعي الثاني (هل تختلف دقة تقدير معلمة تمييز الفقرة باستخدام النموذج الثنائي المعلمة لصور الاختبار الثلاث باختلاف زمن الاختبار؟)

تم حساب المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم تمييز الفقرات لكل صورة من صور الاختبار.

جدول (20):المتوسطات الحسابية والانحرافات المعيارية للخطأ لمعياري لتقدير معلم التمييز في النموذج الثنائي وفقا لمتغير وقت الاختبار

الوقت	العدد	المتوسط الحسابي للخطأ المعياري	الانحراف المعياري
45	30	0.141	0.173
60	29	0.067	0.021
75	30	0.069	0.017
الكلي	89	0.092	0.106

كانت المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم تمييز لفقرت كل صورة (0.141، 0.067،0.069)

حسب الترتيب، وهي قيم تشير إلى دقة تقدير معالم تمييز الفقرات لصور الاختبار الثلاث كما هو موضح في جدول (20).

ولفحص دقة تقدير معالم التمييزفقد تم استخدام تحليل التباين الأحادي وفق تصميم القياسات المتكررة للأخطاء المعيارية لهذه التقديرات لكل من صور الاختبار الثلاث، والموضحة في الجدول (21).

جدول (21): نتائج تحليل التباين الأحادي للخطأ لمعياري لتقدير معلم التمييز في النموذج الثنائي وفقا لمتغير وقت الاختبار

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	ف	الدلالة الإحصائية
وقت الاختبار	0.105	2	0.052	5.102	0.008
الخطأ	0.884	86	0.010		
الكلي	0.988	88			

ويلاحظ من الجدول (21) وجود فرق ذات دلالة إحصائية في تقدير تمييز الفقرات تعزى إلى صور الاختبار، حيث كان متوسط الأخطاء المعيارية للصورة الأولى الثانية والثالثة أعلى منه للصورتين الثانية والثالثة، ولذلك تم استخدام اختبار المقارنات البعدية بين متوسطات الأخطاء المعيارية لتقدير معلم التمييز لفقرات الصور الثلاث للاختبار لمعرفة أي الفروق دالة إحصائيا.

ويبين الجدول (22) نتائج المقارنات الثنائية بين متوسطات الأخطاء المعيارية.

جدول (22): المقارنات البعدية بين المتوسطات الحسابية للخطأ لمعياري لتقدير معلم التمييز في النموذج الثنائي وفقا لمتغير وقت الاختبار

المقارنات	الفرق بين المتوسطات	Std. Error	الدلالة الإحصائية
45 * 60	0.074	0.026	0.024
45* 75	0.072	0.026	0.028
60 * 75	-0.002	0.026	0.997

ويلاحظ من الجدول (22) أن الفرق بين متوسطي الأخطاء المعيارية كان دالا إحصائيا بين الصورتين الأولى، والثالثة، ولصالح الصورة الأولى؛ أي أن فقرات الصورة الثالثة للاختبار أكثر دقة في تقديرات معالم تمييز من فقرات الصورة الأولى للاختبار. كما أشارت إلى أن الفرق بين متوسطي الأخطاء المعيارية كان دالا إحصائيا بين الصورتين الأولى والثانية، ولصالح الصورة الأولى؛ أي أن فقرات الصورة الثانية للاختبار أكثر دقة في تقديرات معالم تمييز من فقرات الصورة الأولى للاختبار. كما أشارت النتائج إلى أنه لا يوجد دلالة إحصائية للفرق بين متوسطي الأخطاء المعيارية بين الصورتين الثانية والثالثة؛ أي أن تقديرات معالم التمييز لا تتأثر بطول زمن الاختبار سواء كان (60) دقيقة أو (75) دقيقة.

خامسا: النتائج المتعلقة بالسؤال الفرعي الثالث (هل تختلف دقة تقدير معلمة قدرة الفرد باستخدام النموذج الأحادي المعلمة (نموذج راش) لصور الاختبار الثلاث باختلاف زمن الاختبار؟)

تم حساب المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم القدرة لأفراد كل صورة من صور الاختبار والموضحة في الجدول (23).

جدول (23):المتوسطات الحسابية والانحرافات المعيارية للخطأ لمعياري لتقدير معلم القدرة في النموذج الأحادي وفقا لمتغير وقت الاختبار

الوقت	العدد	المتوسط الحسابي للخطأ المعياري	الانحراف المعياري
45	476	0.482	0.028
60	486	0.723	0.010
75	498	0.638	0.026
الكلي	1460	0.615	0.102

ولفحص دقة تقدير معالم القدرة، فقد تم استخدام تحليل التباين الأحـــادي وفـــق تـــصميم القياسات المتكررة للأخطاء المعيارية لهذه التقديرات لكل من صور الاختبار الثلاث، والموضحة في الجدول (24).

جدول (24): نتائج تحليل التباين الأحادي للخطأ لمعياري لتقدير معلم القدرة في النموذج الأحادي وفقا لمتغير وقت الاختبار

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	ف	الدلالة الإحصائية
وقت الاختبار	14.280	2	7.140	13851.499	.001
الخطأ	.751	1457	.001		
الكلي	15.031	1459			

ويلاحظ من الجدول (24) إلى وجود فرق ذا دلالة إحصائية في تقدير صعوبة الفقرات يعزى إلى صور الاختبار، حيث كان متوسط الأخطاء المعيارية للصورة الثانية أعلى منه للصورتين الأولى والثالثة.

ولذلك تم استخدام طريقة المقارنات البعدية بين متوسطات الأخطاء المعيارية لتقدير معلمة القدرة للأفراد للصور الثلاث للاختبار لمعرفة أي الفرق دالة إحصائيا. ويبين الجدول (25) نتائج المقارنات الثنائية بين متوسطات الأخطاء المعيارية.

(25): نتائج المقارنات البعدية لمتوسطات الخطأ لمعياري لتقدير معلم القدرة في النموذج الأحادي وفقا لمتغير وقت الاختبار

المقارنات	الفرق بين المتوسطات	Std. Error	الدلالة الإحصائية
45 * 60	-0.240	0.001	.001
45* 75	-0.156	0.001	.001
60 * 75	0.085	0.001	.001

ويلاحظ من الجدول (25) أن الفرق في متوسطي الأخطاء المعيارية كان دالا إحصائيا بين الصورتين الأولى والثانية ولصالح الثانية أي أن فقرات الصورة أكثر دقة في تقديرات القدرة من فقرات الصورة الثانية. كما أشارت إلى أن الفرق بين متوسطي الأخطاء المعيارية كان دالا إحصائيا بين الصورتين الأولى والثالثة ولصالح الصورة الثالثة؛ أي أن فقرات الصورة الأولى للاختبار أكثر دقة في تقديرات معالم القدرة من فقرات الصورة الثالثة للاختبار. كما أشارت

النتائج إلى أنه يوجد دلالة إحصائية للفرق في متوسطات الأخطاء المعيارية بين الصورتين الثانية والثالثة لصالح الصورة الثانية؛ أي أن تقديرات معالم القدرة تتأثر بطول زمن الاختبار عندما يكون (75) دقيقة أكثر منه عندما يكون الزمن (60) دقيقة.

النتائج المتعلقة بالسؤال الفرعي الثالث (هل تختلف دقة تقدير معلمة قدرة الفرد باستخدام النموذج الثنائي المعالم لصور الاختبار الثلاث باختلاف زمن الاختبار؟)

تم حساب المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم القدرة لأفراد كل صورة من صور الاختبار والموضحة في الجدول (26).

جدول (26): المتوسطات الحسابية والانحرافات المعيارية للخطأ لمعياري لتقدير معلم القدرة في النموذج الثنائي وفقا لمتغير وقت الاختبار

الوقت	العدد	المتوسط الحسابي للخطأ المعياري	الانحراف المعياري
45	476	0.303	0.228
60	486	0.550	0.038
75	498	0.537	0.048
الكلي	1460	0.465	0.176

وللتحقق من أثر زمن الاختبار تم إجراء تحليل التباين الأحادي لمتوسطات الأخطاء المعيارية في تقدير هذه المعالم باستخدام برنامج (SPSS) كما في الجدول (27).

جدول (27): نتائج تحليل التباين الأحادي للخطأ المعياري لتقدير معلم القدرة في النموذج الثنائي وفقا لمتغير وقت الاختبار

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	ف	الدلالة الإحصائية
وقت الاختبار	18.621	2	9.311	510.993	.001
الخطأ	26.547	1457	.018		
الكلي	45.168	1459			

ويلاحظ من الجدول وجود فرق ذا دلالة إحصائية في تقدير قدرة الأفراد تعزى إلى صور الاختبار، حيث كان متوسط الأخطاء المعيارية للصورة الثانية أعلى منه للصورة الأولى والثالثة. ولذلك تم استخدام طريقة المقارنات البعدية بين متوسطات الأخطاء المعيارية لتقدير قدرة الأفـراد للصور الثلاث للاختبار لمعرفة أي الفروق دالة إحصائيا. ويبين الجدول (28) نتائج المقارنـات الثنائية بين متوسطات الأخطاء المعيارية.

جدول (28): المقارنات البعدية بين متوسطات الخطأ المعياري لتقدير معلم القدرة في النموذج الثنائي وفقا لمتغير وقت الاختبار

المقارنات	الفرق بين المتوسطات	Std. Error	الدلالة الإحصائية
45 * 60	-0.247	0.009	.001
45* 75	-0.234	0.009	.001
60 * 75	0.013	0.009	.348

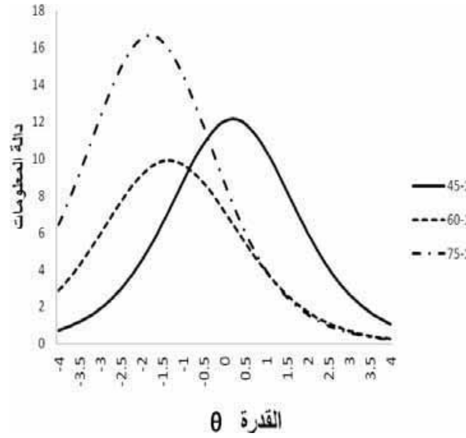
ويشير الجدول (28) إلى أن الفرق بين متوسطي الأخطاء المعيارية كانت دالا إحصائيا بين الصورتين الأولى والثانية ولصالح الثانية؛ أي أن فقرات الصورة الأولى أكثر دقة في تقديرات القدرة من فقرات الصورة الثانية. كما أشارت إلى أن الفرق بين متوسطي الأخطاء المعيارية كان دالا إحصائيا بين الصورتين الأولى والثالثة ولصالح الصورة الثالثة؛ أي أن فقرات الصورة الأولى للاختبار أكثر دقة في تقديرات معالم القدرة من فقرات الصورة الثالثة للاختبار. كما أشارت النتائج إلى أنه لا يوجد دلالة إحصائية للفرق بين متوسطي الأخطاء المعيارية بين الصورتين الثانية والثالثة؛ أي أن تقديرات معالم قدرة الأفراد لا تتأثر بطول زمن الاختبار سواء كان (60) دقيقة أو (75) دقيقة.

سادسا: النتائج المتعلقة بقيم دالة المعلومات للفقرة والاختبار لصور الاختبار الثلاث

النتائج المتعلقة بالسؤال الفرعي الرابع (هل تختلف كمية المعلومات للفقرات وللاختبار عند كل مستوى من مستويات القدرة لأفراد اختبار اختيار من متعدد باختلاف الزمن المعطى للاختبار والنموذج المستخدم (أحادي المعلمة، ثنائي المعلمة))؟

النتائج المتعلقة بقيم دالة المعلومات للفقرات ودالة المعلومات للاختبار بصوره الثلاث وفق النموذج الأحادي (نموذج راش)

لإيجاد دالة المعلومات للفقرات ودالة المعلومات للاختبار لصور الاختبار الثلاث وفق النموذج الأحادي (نموذج راش) تم استخدام البرنامج الإحصائي (BILOG – MG) المصمم لتحليل البيانات وفق نماذج نظرية استجابة الفقرة، ويوضح الشكل (4) منحنيات دالة المعلومات لصور الاختبار الثلاث التي توضح كمية المعلومات التي يقدمها الاختبار عند مستويات القدرة المختلفة.

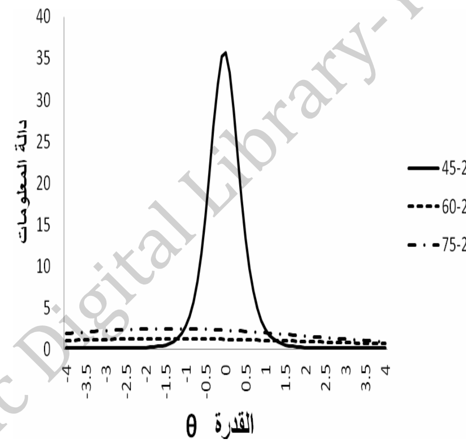


الشكل (4): دالة المعلومات لصور الاختبار الثلاث باستخدام النموذج الأحادي المعلمة

ويتبين من الشكل (4) أن الصورتين الثانية والثالثة أعطت أكبر قدر من المعلومات عند مستويات القدرة المنخفضة، بينما الصورة الأولى أعطت أكبر قدر من المعلومات عند مستويات القدرة المتوسطة، وأعطت الصور الثلاث أقل كمية من المعلومات عند مستويات القدرة المرتفعة؛ بمعنى أن الصور الثلاث أعطت معلومات قليلة عن الطلبة ذوي القدرات المرتفعة. كما كانت الصورة الثالثة هي الأعلى في كمية المعلومات تليها الصورة الثانية ثم الصورة الأولى عند مستويات القدرة المنخفضة؛ أي أن الصورة الثالثة (إعطاء الطلبة (75) دقيقة لإكمال الاختبار) أعطت معلومات أكثر عن ذوي القدرات المتدنية؛ أي أن الطلبة ذوي القدرة المنخفضة قد استفادوا من تمديد زمن الاختبار. أما عند مستويات القدرة المتوسطة فإن الصورة الثالثة هي الأعلى أيضا تليها الصورة الأولى ثم الصورة الثانية؛ أي أن الصورة الثالثة (إعطاء الطلبة (75) دقيقة لإكمال الاختبار) تعطي معلومات أكثر عن ذوي القدرات المتوسطة؛ أي أن الطلبة ذوي القدرة المتوسطة قد استفادوا من تمديد زمن الاختبار. أما عند مستويات القدرة المرتفعة فإن الصور الثلاث قد أعطت كميات معلومات قليلة ؛ أي أن الطلبة ذوي القدرة المرتفعة لم يستفيدوا من تمديد زمن الاختبار.

النتائج المتعلقة بقيم دالة المعلومات للفقرات والاختبار لصور الاختبار الثلاث وفق النموذج الثنائي المعلمة

لإيجاد اقتران المعلومات للفقرات والاختبار لصور الاختبار الثلاث وفق النموذج الثنائي تم استخدام البرنامج الإحصائي (BILOG–MG) المصمم لتحليل البيانات وفق نماذج نظرية استجابة الفقرة. ويوضح الشكل (5) منحنيات دالة المعلومات لصور الاختبار الثلاث التي توضح كمية المعلومات التي يقدمها الاختبار عند مستويات القدرة المختلفة.



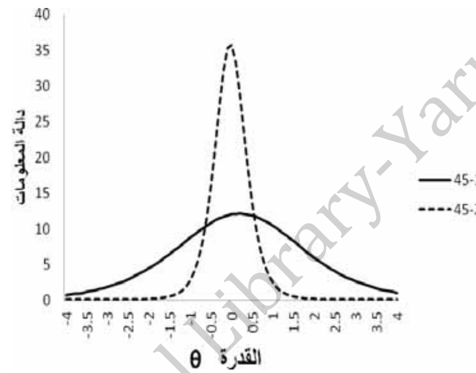
الشكل (5): دالة المعلومات لصور الاختبار الثلاث الاختبار

باستخدام النموذج الثنائي المعالم

ويتبين من الشكل (5) أن الصورة الأولى أعطت أكبر قدر من المعلومات عند مستويات القدرة المتوسطة فقط، وأن الصور الثلاث أعطت أقل كمية عند مستويات القدرة المنخفضة والمرتفعة؛ أي أن الصورة الأولى من الاختبار (إعطاء الطلاب 45 دقيقة لإكمال الاختبار) أعطت

معلومات أكثر عن ذوي القدرة المتوسطة؛ أي أن الطلبة ذوي القدرة المتوسطة لم يتأثروا بنقصان زمن الاختبار.

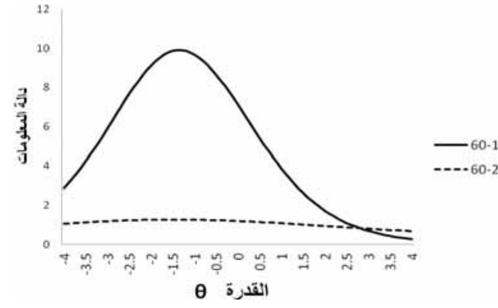
ويوضح الشكل (6) دالة المعلومات لصورة الاختبار الأولى باستخدام النموذج أحـــادي المعلمة ودالة المعلومات لصورة الاختبار الأولى باستخدام النموذج الثنائي المعالم.



الشكل (6): دالة المعلومات لصورة الاختبار الأولى باستخدام النموذجين

يلاحظ من الشكل (6) أن الصورة الأولى للاختبار وفقا للنموذج الأحادي أعطت اكبر قدر ممكن من المعلومات عند مستويات القدرة المتوسطة، و أعطت كمية معلومات قليلة عند مستويات القدرة المنخفضة والمرتفعة. أما بالنسبة للصورة الأولى وفقا للنموذج الثنائي المعالم فأعطت أكبر قدر من المعلومات عند مستويات القدرة المتوسطة فقط.

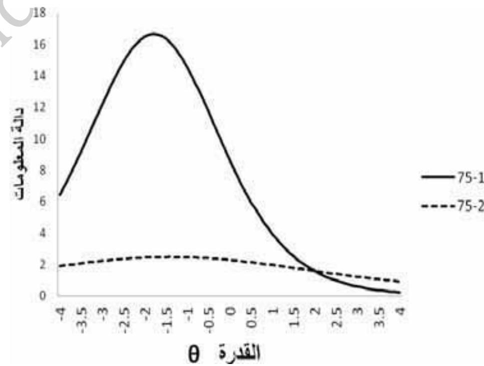
ويبين الشكل (7) دالة المعلومات لصورة الاختبار الثانية باستخدام النموذج أحادي المعلمة ودالة المعلومات لصورة الاختبار الثانية باستخدام النموذج الثنائي المعالم.



الشكل (7): دالة المعلومات لصورة الاختبار الثانية باستخدام النموذجين

أما بالنسبة لدالة المعلومات للصورة الثانية، فقد كانت كمية المعلومات وفقا للنموذج الأحادي المعلمة أعلى منها لكمية المعلومات وفقا للنموذج الثنائي لمستويات القدرة المتدنية وكذلك الحال بالنسبة لمستويات القدرة المتوسطة والعالية كما هو موضح في الشكل (7).

ويبين الشكل (8) دالة المعلومات لصورة الاختبار الثانية باستخدام النموذج أحادي المعلمة ودالة المعلومات لصورة الاختبار الثانية باستخدام النموذج الثنائي المعالم.



الشكل (8): دالة المعلومات لصورة الاختبار الثالثة باستخدام النموذجين

أما بالنسبة للصورة الثالثة، فقد كانت كمية المعلومات وفقا للنموذج الأحادي أعلى منها لكمية المعلومات وفقا للنموذج الثنائي المعالم لمستويات القدرة المختلفة كما هو واضح قي الشكل (8).

الاستنتاجات والتوصيات

في ضوء نتائج هذه الدراسة تبين أن هنالك بعض الاستنتاجات التي لا بد من الإشارة إليها، وأن هناك بعض التوصيات التي يمكن الاستفادة منها، مع التأكيد على أن حال هذه الدراسة كحال معظم الدراسات التي سبقت وناقشت مثل هذه الطرق، ونتائج دراسة واحدة قد لا تؤكد تفوق أسلوب أو طريقة بشكل مطلق، وخصوصا أن هذا المجال ما زال بحاجة إلى الكثير من البحث والتقصي.

أظهرت نتائج التحليل العاملي للاختبار بصورتيه الثانية والثالثة باستخدام المكونات الرئيسية أن قيم التباين المفسر للعامل الأول كانت عالية للصورتين بالمقارنة مع ما يفسره العامل الثاني لتلك الصور من الاختبار. وهذا مؤشر على تحقق افتراض أحادية البعد، وفق المؤشرات المعتمدة على المكونات الرئيسية، حيث أشار ريكاس (Reckase,1979) إلى أنه إذا فسر العامل الأول أكثر من 20%، فإن هذا دليل على أحادية البعد. وأظهرت نتائج التحليل العاملي لصورتي الاختبار الثانية والثالثة كذلك أن قيم الجذر الكامن الأول إلى الجذر الكامن الثاني كانت أعلى من 2. وهذا مؤشر آخر على تحقق افتراض أحادية البعد بالاعتماد على اليستير وهوتشينسين (Alastair & Hutchinsen, 1987).

وأكد استخدام ما يعرف ب Scree Plot تحقق افتراض أحادية البعد في الصورتين الثانية والثالثة من الاختبار،حيث تبين أن الجذر الكامن للعامل الأول تميز بشكل واضح عن الجذور الكامنة لبقية العوامل، وذلك لأن نسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني كانت عالية.

ولكن أظهرت نتائج التحليل العاملي للاختبار بصورته الأولى باستخدام المكونات الرئيسية أن قيمة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني كانت أقل من 2، وهذا مؤشر على عدم تحقق افتراض أحادية البعد،وظهور عامل آخر يؤثر على قياس السمة المراد قياسها قد يكون عامل السرعة،حيث أشار مروخ (Mroch,2006) إلى أن انتهاك افتراض التحرر من السرعة يؤدي إلى انتهاك افتراض أحادية البعد في نظرية استجابة الفقرة، وذلك لأن سرعة الاختبار تضيف بعدا إحصائيا آخر لما يقيسه الاختبار؛ أي أن هنالك عامل آخر في الاختبارات السريعة يمكن استخراجه، وبالتالي يصبح الاختبار السريع متعدد الأبعاد.

وعند مطابقة الفقرات للنماذج اللوجستية المستخدمة باستخدام البرنامج الإحصائي (BILOG – MG) تبين من عدد الفقرات المطابقة لكل صورة من صور الاختبار الثلاث أن فقرات الصورة الثالثة للاختبار كانت أكثر مطابقة مع النماذج اللوجستية المستخدمة في الدراسة (النموذج فقرات اللوجستي الأحادي المعلمة والنموذج اللوجستي ثنائي المعلمة) ؛ أي أن عند إعطاء الطلاب زمنا أكثر(أي 75 دقيقة) فانه يحسن من مطابقة الفقرات للنماذج اللوجستية المستخدمة. ويستنتج بأنه قد يستخدم الاختلاف في عدد الفقرات المطابقة لتقديم مؤشرات على تأثير زمن الاختبار على المطابقة للنموذج.

وأظهرت النتائج أن التقديرات لمعالم صعوبة فقرات الصورة الأولى للاختبار أكثر دقة منها لمعالم صعوبة فقرات الصورة الثانية والثالثة، وكانت التقديرات لمعالم صعوبة فقرات الصورة الثالثة للاختبار أكثر دقة منها لمعالم صعوبة فقرات الصورة الثانية للاختبار وفق النموذج الأحادي المعلمة (نموذج راش). وهذا يعني تأثر دقة تقديرات معالم صعوبة الفقرة بزمن الاختبار؛أي أن السرعة (المتمثلة بزمن الاختبار) تؤثر على دقة تقديرات معالم الصعوبة ،وهذا يتفق ونتائج دراسة أوشيما (Oshima,1994).

ووفق النموذج الثنائي المعالم، كانت التقديرات لمعالم صعوبة فقرات الصورة الأولى للاختبار أكثر دقة منها لمعالم الصعوبة لفقرات الصورة الثانية والثالثة، وكانت التقديرات لمعالم صعوبة فقرات الصورة الثانية للاختبار أكثر دقة منها لمعالم الصعوبة لفقرات الصورة الثالثة. وبالتالي استنتج الباحث بأن دقة تقديرات صعوبة الفقرة تختلف باختلاف صور الاختبار؛ (أي باختلاف زمن الاختبار) أي أن السرعة (المتمثلة بزمن الاختبار) تؤثر على دقة تقديرات معالم الصعوبة، وهذا يتفق ونتائج دراسة أوشيما (Oshima,1994).

كما أظهرت النتائج أن التقديرات لمعالم تمييز فقرات الصورة الثانية أكثر دقة منها لمعالم تمييز فقرات الصورة الأولى، كما أشارت إلى أن دقة تقديرات معالم التمييز للصورة الثالثة أكثر دقة منها لتقديرات معالم التمييز للصورة الأولى للاختبار، مع عدم وجود دلالة إحصائية بين متوسطات الأخطاء المعيارية في تقديرات معالم التمييز للفقرات بين الصورتين الثانية والثالثة من الاختبار، وذلك وفق النموذج الثنائي المعلمة. وبالتالي استنتج الباحث بأن دقة تقديرات تمييز الفقرة تختلف باختلاف صور الاختبار؛ (أي باختلاف زمن الاختبار) أي أن السرعة (المتمثلة

بزمن الاختبار) تؤثر على دقة تقديرات معالم الصعوبة ،وهذا يتفق ونتائج دراسة أوشيما (Oshima,1994).

وأظهرت النتائج أيضا أن التقديرات كانت أكثر دقة لقدرات أفراد الصورة الأولى للاختبار منها لتقديرات قدرات أفراد الصورة الثانية والثالثة له، وكانت التقديرات لقدرات أفراد الصورة الثالثة أكثر دقة منها للصورة الثانية وفق النموذج الأحادي المعلمة (نموذج راش). ووفق النموذج الثنائي المعلمة كانت التقديرات أكثر دقة لقدرات أفراد الصورة الأولى للاختبار منها لتقديرات قدرات أفراد الصورة الثانية والثالثة له، مع عدم وجود دلالة إحصائية بين متوسطات الأخطاء المعيارية لتقديرات قدرات أفراد الصورتين الثانية والثالثة من الاختبار. وبالتالي أستنتج الباحث بأن دقة تقديرات معاملات قدرة الفرد تختلف باختلاف صور الاختبار؛ (أي باختلاف زمن الاختبار) أي أن السرعة (المتمثلة بزمن الاختبار) تؤثر على دقة تقديرات معالم الصعوبة، وهذا يتفق ونتائج دراسة أوشيما (Oshima,1994).

وعند فحص دالة المعلومات لصور الاختبار الثلاث باستخدام النموذج اللوجستي الأحادي المعلمة (نموذج راش)، وجد أن الصورة الأولى أعطت أكبر قدر من المعلومات عند مستويات القدرة المتوسطة، وأن الصورة الثالثة والثانية أعطت أكبر قدر من المعلومات عند مستويات القدرة المنخفضة؛ وهذا يعني أن الطلبة ذوي القدرة المنخفضة والمتوسطة فقط قد استفادوا من تمديد وقت الاختبار. أما عند فحص دالة المعلومات لصور الاختبار الثلاث باستخدام النموذج اللوجستي الثنائي المعالم تبين أن الصورة الأولى كانت اكبر كمية معلومات عند مستويات

القدرة المتوسطة فقط؛ وهذا يعني أن الطلبة ذوي القدرة المتوسطة لم يتأثروا من تمديد زمن الاختبار.

وبذلك، فإن عامل السرعة عامل مؤثر بشكل جوهري في جميع التحليلات الخاصة بمعالم الفقرات ومعالم الأفراد ودقة التقديرات وفق البرامج المستخدمة مثل BILOG – MG، ومن هنا تأتي أهمية إيجاد التوازن بين عدد الفقرات في الاختبار وزمن الإجابة وفقا لخصائص المفحوصين حسب قدراتهم (العالية والمتوسطة والمنخفضة)، أو الزمن الذي يحقق أعلى توازن بوجود الفئات الثلاث مجتمعة في موقف اختباري واحد، وتشير النتائج إلى أن افتراض التحرر من السرعة لا يتحقق بسهولة، فهو عامل مؤثر في تحقق الافتراضات الأخرى، فعندما تفترض النظرية أحادية البعد، فإن التحرر من السرعة ليس افتراضا آخرا بل هو الافتراض الذي يشكل شرطا مسبقا لتحقق أحادية البعد، ولكن ليس شرطا كافيا، فالأحادي البعد لا بد وان يكون متحررا من السرعة، أما المتحرر من السرعة ليس شرطا أن يكون أحادي البعد، أما الأهم في الآثار الجانبية لعامل السرعة هو دقة القرارات أو الثبات في تفسير نتائج الاختبار، وقد لاحظنا تأثيره على مقدار المعلومات (Information) وكذلك على صدق المحتوى من خلال نسبة الفقرات المتبقية في ضوء درجة مطابقتها لنماذج نظرية استجابة الفقرة.

وفي ضوء تحليل التباين الأحادي للخطأ المعياري تبين بأنه لا توجد صورة واحدة من صور الاختبار كانت هي الأكثر دقة في تقدير معالم الفقرة فمن هنا توصي الدراسة بان **يستخدم أسلوب ثالث يتضمن إعادة تحليل البيانات بعد استبعاد الفقرات الغير مطابقة ثم إجراء التحليل من جديد على الفقرات المتبقية لعلها تخرج صورة من صور الاختبار تكون الأكثر دقة في تقدير**

معالم الفقرة والقدرة. وكما لوحظ من خلال هذه الدراسة بأنه تم استخدام نمـــوذجين لوجـــستيين، النموذج اللوجستي الأحادي المعلمة، والنموذج اللوجستي الثنائي المعالم **ربمـــا إذا تـــم اســـتخدام النموذج الثلاثي المعالم أن توجد صورة واحدة من صور الاختبار تكون أكثر دقة في تقدير معالم الفقرة والقدرة.**

المراجع.

البخاتي، فلاح. (2009). بناء اختبار تحصيلي لمادة مناهج البحث العلمي على وفق نظرية القياس التقليدية وأنموذج راش. رسالة دكتوراه غبر منشورة، جامعة بغداد، كلية التربية، بغداد: العراق.

الشريفين، نضال وطعامنة، إيمان. (2009).اثر عدد البدائل في اختبار الاختيار من متعدد في نظرية الاستجابة للفقرة، المجلة الأردنية في العلوم التربوية، 5، (4)، 309 – 332.

علام، صلاح الدين. (1986). تطورات معاصرة في القياس النفسي التربوي. جامعة الكويت، الكويت: إدارة التأليف والنشر.

علام، صلاح الدين. (2005). نماذج الاستجابة للمفردة أحادية البعد ومتعددة الأبعاد. القاهرة: دار الفكر العربي.

النعيمي، عزالدين. (2006). اثر انتهاك افتراض الاستقلال الموضعي على التقديرات المختلفة لنظرية استجابة الفقرة، رسالة دكتوراه غبر منشورة، جامعة اليرموك، اربد: الأردن.

المراجع الأجنبية

Alastair. P., & Hutchinson, G. (1987). Calibrating graded assessment Rash partial credit analysis of performance in Writing Language testing, 4, 73-92.

Albanese, M. A., & Forsyth, R.A. (1984). The one-two and modified two-parameter latent trait models. An empirical study of relative fit. *Educational and Psychological Measurement, 44,229-246*

Allen,M.J., & Yen, W.M. (1979). Introduction to measurement theory. California: Cole Publishing Company.

Armitage, A.(1999). Extended time limits on university examination. A master thesis. University of Calgray.

Bejar, I.I. (1980). A procedure for investigating the unidimensionality of achievement test based on item parameter estimates. *Journal of Educational Measurement*, 17 (4), 283.

Bejar, I.I. (1985). Test speededness under number-right scoring: An analysis of the test of English as a Foreign Language (Report No.ETS-RR-85-11).Princton, NJ: Educational Testing Service.

Bolt, D.M, Cohen, A.S. & Wollack, J.A. (2002). Item parameter, estimation under conditions of test speededness: Application of a mixture Rash model with ordinal constraints .*Journal of Educational Measurement*, 39, 331-348.

Bridgeman, B., Cline, F., & Hessinger, J. (2003). Effect of extra time on GRE Quantities and Verbal Scores (Research Report 03-13), Princeton NJ: Educational Testing Service.

Bridgeman, B., McBride, A., & Monaghan, W. (2004). Testing and time limits.R&D Connection.Educational Testing Service (ETS).

Briel, J.B., O'Neill, K.A., and Scheuneman, J.D. (1993). GRE technical manual. Princeton, NJ: Educational Testing Service.

Crocker, L. & Algina, J. (1986). Introduction to classical and modern test theory. California: COLE Publishing Company.

Cronbach, L.J., & Warrington, W.G. (1951). Time limit tests: Estimating their reliability and degree of speeding. *Psychometrika, 14, 167-188.*

Crone, C., Wright., & Baron, P. (1993) .Performance of examinees for whom English is there second language on the spring 1992 SAT11: Writing test (Unpublished Manuscript). Princton, NJ: Educational Testing Service.

Donlon, T.F. (1973). Establishing time limits for tests. A paper presented at the Annual Meeting of the Northeast Educational Research Association, Ellenville, NY.

Douglas, J., Kim, H.R., Habing, B., & Gao, F. (1998). Investigating local dependence with conditional covariance functions, *Journal of Educational and Behavioral Statistics, 23, 129-151*.

Gulliksen, H. (1950). Theory of mental test. New York, NY: John Wiley.

Hale, G.A. (1992). Effect of amount of time allowed on the test written English (TOEFL Research Report #PR-92-27). Princton, NJ: Educational Testing Service.

Hambleton, R.K., & Swaminathan, H. (1985). Item response theory: Principles and applications. Boston: Kluwer – Nijhoff.

Hambleton, R.K., & Traub, R.R. (1971). Information curves and efficiency of three logistic test models. *British Journal of Mathematical and Statical Psychology.24.273-281.*

Hecht, L.W, Swinford, F., (1981). Item analysis at educational testing serices. Unpublished Statistical Report, Educational Testing Services, Princton, NJ.

Kingston, N.M., & Dorans, N.J. (1984). Item location effects and their implications for IRT equating and adaptive testing. *Applied Psychological Measurement, 8, 147-154.*

Livingston, S.A. (1987). The effect of time limits on the quality of student-written essays. Paper presented at the annual meeting of the American Educational Research Associated, New York, NY.

Lord, F.M. (1980). Application of item response theory to practical testing problems Mahwah. NJ:Laurence Erlbaum Associates.

Lu, Y., & Sireci, S.G. (2007). Validity issues in test speededness. *Educational Measurement, 26(3), 29-37.*

Mandinach, E.B., Bridgeman, B., Cahalan. - Laitusis, C. & Trapani. C. (2005). The impact of extended time on SAT test performance, College Board Research Report no. 2005-8 ETS RR-05-20.

Meara, K., Robin, F.& Sireci, S.(2000). Using Multidimentional scaling to assess the dimentionality of dichotomous item data. *Multivaraite Behavioral Research, 35,229-259*

Mroch. A.A. (2006). A randomized response deletion approach for reducing the effects of test speededness on item response theory item parameter estimates. Doctoral dissertation. University of Wisconsion – Madison.

Nunnally, J.C. (1978). Psychometric Theory. McGraw – Hill: New York.

Oshima, T.C. (1994). The effect of speededness on parameter estimation in item response theory. *Journal of Educational Measurement, 31, 200-219.*

Peterson, Norman G. (1993). Review of issues associated with speededness of GATB tests. American Institutes for Research.

Powers, D.E., & Fowles, M.E. (1996). Effect of applying different time limits to a proposed GRE writing test. *Journal of Educational Measurement, 33 (4), 433-452.*

Reckase, M.D. (1979). Unifactor latent trait models applied to multifactor test.*Journal of Educational Statistics. 4. 207-230.*

Rindler, S.D. (1979). Paitfalls in assessing test speededness. *Journal of Educational Measurement, 16, 261-270.*

Schnipke, D. L. (1995). Assessing speededness in computer – based test using item respond time. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco.

Schnipke, D. L. (1999). The influence of speedeness on item-parameter estimation.Law school admission council: computerized testing report 96-07.

Swinford, F. (1979).The test consultants manual, Educational Testing Services, Princton, NJ.

Wayne, J.C. (March, 2003). Comparing essays written under different timing conditions. Office of Research and Development. The college Board, SAT and Logo, New York.

Zimowski, M.F., Muraki, E.Mislevy, R.J., & Bock, R.D. (2002). BILOG-MG (Version 3) [Computer Software]. Lincolnwood, IL: Scientific Software International 73-92.

الملاحق

ملحق (1): عدد الأهداف في كل مستوى

الوحدة الرابعة : الهندسة الإحداثية

4-1 : المسافة بين نقطتين

الأهداف

1- أن يعرّف الطالب مفهوم الهندسة الإحداثية . **(معرفة)**

2- أن يرسم الطالب محورين متعامدين باستخدام ورقة الرسم البياني .**(فهم)**

3- أن يحدد الطالب إحداثيات نقطة في المستوى البياني . **(فهم)**

4- أن يعين الطالب نقاطا في المستوى البياني .**(فهم)**

5- أن يجد الطالب المسافة بين نقطتين باستخدام نظرية فيثاغورس . **(تطبيق)**

6- أن يستنتج الطالب قانون المسافة بين نقطتين . **(تحليل)**

7- أن يجد الطالب المسافة بين نقطتين باستخدام القانون . **(تطبيق)**

8- أن يطبق الطالب قانون المسافة بين نقطتين لحل مسائل كلامية. **(تطبيق)**

4-2 : إحداثيات نقطة منتصف قطعة مستقيمة

الأهداف

9- أن يستنتج الطالب قانون إحداثيات نقطة منتصف القطعة المستقيمة . **(تحليل)**

10- أن يجد الطالب منتصف قطعة مستقيمة باستخدام القانون. **(تطبيق)**

11- أن يطبق الطالب قانون إحداثيات نقطة منتصف القطعة المستقيمة لحل مسائل كلامية . **(تطبيق)**

4-3 : معادلة الخط المستقيم

الأهداف

12- أن يعرّف الطالب قاعدة ميل الخط المستقيم . **(معرفة)**

13- أن يجد الطالب ميل الخط المستقيم المار بنقطتين . **(تطبيق)**

14- أن يستنتج الطالب معادلة الخط المستقيم الذي ميله (م) ويمر بالنقطة (س1،ص1) .

(تحليل)

15- أن يجد الطالب معادلة الخط المستقيم الذي علم منه ميله وإحداثيات نقطة يمر فيها.

(تطبيق)

16- أن يجد الطالب معادلة الخط المستقيم الذي يمر بنقطتين معلوم إحداثياتهما . **(تطبيق)**

17- أن يبين الطالب أي النقاط المعطاة لديه تمر بمستقيم معطى له معادلته . **(فهم)**

18- أن يجد الطالب معادلة الخط المستقيم المعلوم ميله ويمر بنقطة تقاطع مستقيمين معلوم معادلتهما . **(تحليل)**

19- أن يجد الطالب معادلة الخط المستقيم إذا عرف الميل والمقطع السيني . **(تطبيق)**

20- أن يجد الطالب معادلة الخط المستقيم إذا عرف الميل والمقطع الصادي . **(تطبيق)**

21- أن يجد الطالب معادلة الخط المستقيم إذا عرف منه المقطع السيني والمقطع الصادي . **(تطبيق)**

4-4 : معادلة الدائرة

الأهداف

22 – أن يوضح الطالب مفهوم الدائرة .**(فهم)**

23- أن يستنتج الطالب معادلة الدائرة التي مركزها نقطة الأصل (0،0) وطول نصف قطرها (ر) هي $س^2 + ص^2 = ر^2$. **(تحليل)**

24- أن يجد الطالب معادلة الدائرة التي مركزها نقطة الأصل (0،0) ومعلوم نصف قطرها . **(تطبيق)**

25- أن يستنتج الطالب معادلة الدائرة التي مركزها النقطة ل (د،هـ) وطول نصف قطرها (ر) هي : $(س - د)^2 + (ص - هـ)^2 = ر^2$. **(تحليل)**

26- أن يجد الطالب معادلة الدائرة إذا عرف إحداثيات مركزها وطول نصف قطرها . **(تطبيق)**

27-أن يجد الطالب مركز الدائرة و نصف القطر لمعادلة دائرة . **(تطبيق)**

28-أن يطبق الطالب على الصيغة العامة لمعادلة الدائرة . **(تطبيق)**

الوحدة الخامسة : النسب المثلثية

5-1 : جيب الزاوية الحادة

الأهداف

1- أن يعرّف الطالب تشابه المثلثات .**(معرفة)**

2- أن يعرّف الطالب نظرية فيثاغورس . **(معرفة)**

3- أن يعرّف الطالب الضلع المقابل للزاوية والوتر في مثلث قائم الزاوية . **(معرفة)**

4- أن يجد الطالب جيب الزاوية الحادة .**(تطبيق)**

5- أن يجد الطالب جيب الزاوية الحادة من القانون .**(تطبيق)**

6- أن يستخدم الطالب الآلة الحاسبة في إيجاد جيب الزاوية .**(تطبيق)**

7- أن يوظف الطالب جيب الزاوية في مسائل عملية . **(تطبيق)**

5-2 : جيب تمام الزاوية الحادة

الأهداف :

8- أن يعرّف الطالب الضلع المجاور للزاوية .**(معرفة)**

9- أن يعرّف الطالب قاعدة جيب تمام الزاوية .**(معرفة)**

10- أن يجد الطالب جيب تمام زاوية حادة في مثلث قائم الزاوية . **(تطبيق)**

11- أن يستخدم الطالب الآلة الحاسبة في إيجاد جيب تمام الزاوية . **(تطبيق)**

12- أن يوظف الطالب جيب تمام الزاوية في مسائل عملية . **(تطبيق)**

5-3 : ظل الزاوية الحادة

الأهداف

13- أن يعرّف الطالب الضلع المقابل والضلع المجاور للزاوية . **(معرفة)**

14- أن يعرّف الطالب قاعدة الظل. **(معرفة)**

15- أن يجد الطالب ظل زاوية حادة في مثلث قائم الزاوية . **(تطبيق)**

16- أن يستخدم الطالب الآلة الحاسبة في إيجاد ظل زاوية معطاة . **(تطبيق)**

17- أن يوظف الطالب الظل في حل المشكلات (المسائل) الحياتية . **(فهم)**

5-4 : العلاقة بين النسب المثلثية

الأهداف

18- أن يعرّف الطالب جيب وجيب تمام وظل الزوايا 30 درجة و45 درجة و60 درجة . **(معرفة)**

19- أن يعرّف الطالب جيب تمام وظل الزاوية الحادة . **(معرفة)**

20- أن يعرّف الطالب النسب المثلثية للزاوية الحادة ومتممتها . **(معرفة)**

21- ان يجد الطالب النسب المثلثية للزاوية الحادة ومتممتها . **(تطبيق)**

22- أن يعرّف الطالب العلاقة بين الجيب وجيب التمام وظل الزاوية. **(معرفة)**

23- أن يجد الطالب العلاقة بين النسب المثلثية.**(تطبيق)**

5-5 : حل المثلث القائم الزاوية

الأهداف

24- أن يحل الطالب المثلث القائم الزاوية . (**تطبيق)**

25- أن يستخدم الطالب الآلة الحاسبة في حساب النسب المثلثية وقياس الزاوية إذا علمت إحدى نسبها المثلثية . (**تطبيق)**

6-5 : زوايا الارتفاع والانخفاض

الأهداف

26- أن يجد الطالب النسب المثلثية للزوايا الحادة . (**تطبيق)**

27- أن يميز الطالب بين زاوية الارتفاع وزاوية الانخفاض والعلاقة بينهما . **(تحليل)**

28- أن يحل الطالب على زوايا الارتفاع والانخفاض . **(تطبيق)**

29- أن يوظف النسب المثلثية في حل المسائل . (**فهم)**

ملحق(2): عدد الأهداف في كل مستوى

المحتوى	معرفة (تذكر)	فهم (استيعاب)	تطبيق	تحليل	المجموع
الوحدة الرابعة الهندسة الإحداثية	2	5	15	6	28
الوحدة الخامسة النسب المثلثية	11	2	15	1	29
المجموع	13	7	30	7	57

ملحق(3): عدد الأسئلة في كل مستوى

المحتوى	معرفة (تذكر) (23%)	فهم (استيعاب) (12%)	تطبيق (53%)	تحليل (12%)	المجموع (100%)
الوحدة الرابعة 41% الهندسة الإحداثية	3	1	7	1	12
الوحدة الخامسة 59% النسب المثلثية	4	2	10	2	18
المجموع100%	7	3	17	3	30

بسم الله الرحمن الرحيم

ملحق (4) : أسئلة الاختبار

المبحث : الرياضيات **مدة الاختبار :**

الصف : التاسع الأساسي **اليوم والتاريخ :**

الاســم : **المدرسة :**

- **ضع دائرة حول الحرف الممثل للإجابة الصحيحة :**

1) ميل الخط المستقيم المار بالنقطتين أ (2 ، −1)، ب (−3 ، 0) يساوي:

أ) −5 ب) $\frac{-1}{5}$ ج) $\frac{-3}{2}$ د) صفر

2) إحدى النقاط التالية تقع على الخط المستقيم الذي معادلته: ص = 3س −1

أ) (0، 1) ب) (0، −1) ج) (1، −2) د) (4، 2)

3) المسافة بين النقطتين أ (0، 2)، ب (5، 2) هي:

أ) (5) وحدات ب) (25) وحدة ج) (10) وحدات د) (9) وحدات

4) إحداثيات منتصف القطعة المستقيمة أ ب حيث أ (1، 2)، ب (3، 0) هي:

أ) (4، 2) ب) (4 ، 1) ج) (2، 2) د) (2، 1)

5) معادلة الخط المستقيم المار بالنقطتين أ (4، 0) ، ب (5، 4) هي:

أ) ص = 4س−16

ب) ص = 2س−5

ج) ص = 2 س+3

د) ص = 2س−3

6) معادلة الخط المستقيم الذي ميله يساوي 3 ويمر بالنقطة (−1، 1) هي:

أ) ص = 2س−4
ب) ص = 3س−5
ج) ص = 3س−4
د) ص = 3س+ 4

7) معادلة الخط المستقيم الذي مقطعه السيني يساوي 3 ومقطعه الصادي يساوي 1 هي:

أ) ص = $\frac{1}{3}$ س+3 ج) ص = س +3

ب) ص = $\frac{-1}{3}$ س +1 د) ص = س + 1

8) معادلة الدائرة التي مركزها م (3، 2) وتمر بالنقطة (6 ، 6) هي:

أ) $(س-3)^2 + (ص - 2)^2 = 25$
ب) $(س-6)^2 + (ص-6)^2 = 5$
ج) $(س-2)^2 + (ص - 3)^2 = 5$
د) $(س-3)^2 + (ص-2)^2 = 5$

9) إحداثيات مركز الدائرة التي معادلتها $س^2 + ص^2$+8س + 4ص −5=0 هي:

أ) (2، 4) ب) (−4، −2) ج) (−2، −4) د) (−5، 4)

10) معادلة الدائرة التي مركزها يقع على المستقيم ص=2 وتمس محور السينات عند النقطة (1، 0) هي:

أ) (س − 1)^2 + (ص − 2)^2 = 4

ب) (س−1)^2 + ص^2 = 2

ج) (س−1)^2 + ص^2 = 1

د) س^2 + (ص−2)^2 = 4

11) إحدى هذه المعادلات تمثل معادلة دائرة:

أ) س^2 + ص^2 −2س −3ص +15 = 0 ج) س^2 + ص^2 − س +ص +4 = 0

ب) س^2 + ص^2 −4س −6ص +2 = 0 د) س^2 + ص^2 − 2س + 3س +7 = 0

12) إحدى النقاط الآتية تقع على الدائرة التي معادلتها
س^2 + ص^2 +4س − 9ص +14 = 0

أ) (5، 7) ب) (7، 5) ج) (7، 0) د) (0، 7)

13) في المثلث أ ب ج القائم الزاوية في ب إن جيب الزاوية أ ج ب تساوي:

أ) $\frac{\text{أ ب}}{\text{أ ج}}$ ب) $\frac{\text{ب ج}}{\text{أ ج}}$ ج) $\frac{\text{أ ب}}{\text{ب ج}}$ د) $\frac{\text{ب ج}}{\text{أ ب}}$

14) جا 30° تساوي

أ) $\frac{2}{3}$ ب) $\frac{1}{2}$ ج) $\frac{1}{\sqrt{2}}$ د) $\frac{\sqrt{3}}{2}$

15) إذا كانت س: زاوية حادة فإن ظاس =

أ) $\frac{\text{جاس}}{\text{جتاس}}$

ب) $\frac{1}{\text{جتاس}}$

ج) $\frac{1}{\text{جاس}}$

د) $\frac{\text{جتاس}}{\text{جاس}}$

16) إحدى العبارات الآتية صحيحة:

أ) جتا^2 س + جا^2 س = 1
ب) جاس + جتاس = 1
ج) جتا^2 س + جتا^2 ص =1
د) جتا^2 ص + جا^{2}س =1

17) جتا 55ْ =

أ) جا35ْ ب) جتا 35ْ د) ظا 35ْ ج) جا 55ْ

18) إذا كان جاسْ = 4 جتا سْ ، سْ زاوية حادة فإن ظاسْ =

أ) $\frac{1}{4}$ ب) 4 ج) $\frac{1}{2}$ د) 2

19) إذا كان أ ب ج مثلث قائم الزاوية في ب وفيه أ ب = ب ج = 1سم
فإن جيب الزاوية أ ج ب يساوي:

أ) $\sqrt{2}$ ب) 1 ج) $\frac{1}{\sqrt{3}}$ د) $\frac{1}{\sqrt{2}}$

20) إذا كان أ ب ج مثلث الزاوية قائم في ب وكان فيه أ ب = ب ج = 1 فإن ظاج يساوي:

أ) $\frac{\sqrt{3}}{2}$ ب) 1 ج) $\frac{1}{\sqrt{3}}$ د) $\frac{3}{4}$

21) إذا كان جا (45 + س) = 0.9659 فإن قيمة جتا (45 – س)

أ) 0,341 ب) 1 ج) 0.9659 د) صفر

22) جا^2 30 + جا^2 60 =

أ) $\frac{1}{2}$ ب) $\frac{\sqrt{3}}{2}$ ج) $\frac{2}{\sqrt{3}}$ د) 1

23) إذا علمت أن أشعة الشمس كانت تميل بزاوية مقدارها 60 درجة عن سطح الأرض وكان طول ظل الشجرة يساوي (10) أمتار فإن ارتفاع الشجرة يساوي:

أ) 400 متراً ب) 100 متراً ج) 30 متراً د- $\sqrt{300}$ متراً

24) اذا كانت س زاوية حادة فإن إحدى العبارات التالية صحيحة:
أ) صفر ≤ جا س≤ 1
ب) 0 ≤ ظا س ≤ 1
ج) جتا س > 1
د) جا س > 1

25) إذا رصد أحمد سفينة في البحر من قمة منارة ترتفع 30 متراً بزاوية انخفاض مقدارها 45 درجة فإن بعد السفينة عن قاعدة المنارة هو:

أ) $\frac{30}{\sqrt{3}}$ متراً ب) $\frac{\sqrt{2}}{30}$ متراً ج) 30 متراً د) (1) متراً واحداً

26) إحدى العبارات الآتية صحيحة

أ) جا 60° = 2 جا 30° ج) جا 60° = 2 جتا 30°

ب) جا 60° = 2 جا 30 جتا 30° د) جا 60° = جتا 30° حا 30°

27) إذا كانت جا س = جتا 2س فإن قيمة س =

أ) 45° ب) 60° ج) 30° د) 40°

28) إحدى العبارات الآتية صحيحة:

أ) جا س = جتا س

ب) جاس = جتا (90 – س)

ج) جاس = 2جتاس

د) جتاس = –جاس

29) إذا كان أ ب ج مثلث قائم الزاوية في ب وكان جا ج = 0,5 و أ ج = 12سم فإن:

أ) أب = 5سم ب) ب ج = 18سم ج) أب= 6سم د) ب ج = 10سم

30) إذا كان أ ب ج مثلث قائم الزاوية في ب وكان فيه أ ب = 3سم و ب ج = $3\sqrt{3}$ سم فإن قياس الزاوية ج =

أ) 60° ب) 30° ج) 45° د) 40°

Abstract

Manaisa, Haitham Abdallah. The effect of test time on the estimations accuracy of items and persons parameters based on the models of Item Response Theory (IRT). Dissertation. Yarmouk University 2012. (Prof: Ahmad Sulieman Audeh)

The study investigated the effect of test time on the estimations accuracy of item parameters and persons parameters based on the models of Item Response Theory (IRT). To achieve the study objectives, a multiple choice achievement test consisting of 30 items in ninth grade mathematics was constructed. The test took three different forms in terms of the test time for the students to answer test items: Form One with forty five minutes, form two with sixty minutes, and form three with seventy five minutes. The test forms were applied on a total sample of 1500 male and female students with 500 students assigned to each form. Data obtained for each form of the test were analyzed separately using (BILOG – MG).

The study finding revealed the following: there were statistically significant difference (α =0.05) among the standard error means of difficulty item parameters and such estimations in the item's difficulty parameters were more accurate in form one of the test than in form two, in form one of the test than form three, and form three than form two based on Rash Model. Also there were statistically significant difference (α =0.05) among the standard error means of difficulty item parameter and such estimations in the item's difficulty parameters were more accurate in form one of the test than in form two, in form one of the test than in form three, in form two of the test than in form three based on Two Parameters Logistic Model. Also there were statistically significant difference (α =0.05) among the standard error means of discrimination item parameter and such estimations in the item's

discrimination parameters were more accurate in form two of the test than in form one, in form three of the test than in form one, and there were no statistically difference among standard error between form two of the test and form three based on Two Parameters Logistic Model. And for the estimations of a person's ability based on Rash Model there were statistically significant difference (α =0.05) among the standard error means of ability person parameter and such estimations in the person's ability parameters were more accurate in form one of the test than in form two, in form one of the test than in form three, and in form three than in form two. And for the estimations of a person's ability based on Two Parameters Model there were statistically significant difference (α =0.05) among the standard error means of ability person parameter and such estimations in the person's ability parameters were more accurate in form one of the test than in form two, in form one of the test than in form three, and there were no statistically difference among standard error between form two of the test and form three.

In addition to that, form two and form three of the test yielded more information at the low ability levels, Where form one yielded more information at the medium ability levels based on Rash Model. Based on Two Parameters Logistic Model, form one of the test yielded more information at the medium ability levels only. The study recommended using three parameters logistic model in the analysis of the study to deal with guessing, and more analysis with different models regarding the way of manipulation for missing responses. (Keywords: Multiple Choice Tests, Item's Difficulty, Item's Discrimination, Person’s Ability, Function of Item Information, Function of Test Information. Test Time).